

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 02:02:48 ; Search time 278 Seconds
(without alignments)
7374.060 Million cell updates/sec

Title: US-10-019-495-26
Perfect score: 3694
Sequence: 1 ctggcaggttcttagtgagc.....gtcgtataaaaaaaaaa 3694

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	370.2	10.0	3113	2	US-08-993-228-20
2	57.4	1.6	1140	3	US-03-023-173-4
3	56.2	1.5	23673	4	US-09-773-816-1
4	55.2	1.5	1836	4	US-09-475-515-47
5	55.2	1.5	1944	4	US-09-475-515-46
6	55.2	1.5	2025	4	US-09-475-515-45
7	54.8	1.5	2539	3	US-09-000-016-3
8	54.8	1.5	2539	4	US-09-514-340-3
9	54.8	1.5	2809	3	US-09-000-016-1
10	54.8	1.5	2809	4	US-09-514-340-1
11	53.6	1.5	1836	4	US-09-475-515-41
12	53.6	1.5	1836	4	US-09-475-515-44
13	53.6	1.5	1944	4	US-09-475-515-37
14	53.6	1.5	1944	4	US-09-475-515-38
15	53.6	1.5	1944	4	US-09-475-515-40
16	53.6	1.5	1944	4	US-09-475-515-43
17	53.6	1.5	2025	4	US-09-475-515-36
18	53.6	1.5	2025	4	US-09-475-515-39
19	53.6	1.5	2025	4	US-09-475-515-42
20	53.6	1.5	2358	4	US-09-475-515-50
21	53.6	1.5	2466	4	US-09-475-515-49
22	53.6	1.5	2547	4	US-09-475-515-48
23	53.6	1.5	4608	4	US-09-475-515-76
24	53.6	1.5	4689	4	US-09-475-515-74
25	53.2	1.4	2205	4	US-09-252-991A-11217
26	53.2	1.4	2454	4	US-09-252-991A-11038
27	52.6	1.4	1308	4	US-09-501-115-31

28 51.6 1.4 909 4 US-09-252-991A-13568 Sequence 13568, A
 C 29 51.6 1.4 3129 4 US-09-252-991A-13873 Sequence 13873, A
 C 30 51.6 1.4 4188 4 US-09-252-991A-13774 Sequence 13774, A
 31 51.6 1.4 8211 4 US-09-252-991A-13656 Sequence 13656, A
 32 50.4 1.4 1317 4 US-09-252-991A-6463 Sequence 6463, Ap
 33 50.2 1.4 1416 4 US-09-252-991A-2955 Sequence 2955, Ap
 C 34 50.2 1.4 1512 4 US-09-252-991A-2728 Sequence 2728, Ap
 C 35 50.2 1.4 7218 1 US-08-232-463-14 Sequence 14, Appl
 C 36 49.8 1.3 2435 4 US-09-484-970B-134 Sequence 134, Appl
 37 49.8 1.3 152331 3 US-09-128-155-16 Sequence 16, Appl
 38 49.4 1.3 3177 1 US-08-042-747A-4 Sequence 4, Appl
 C 39 49.4 1.3 364 4 US-09-621-976-17202 Sequence 17202, A
 C 40 48.8 1.3 408 4 US-09-252-991A-3047 Sequence 3047, A
 C 41 48.8 1.3 1752 4 US-09-679-686B-17 Sequence 17, Appl
 42 48.8 1.3 13842 3 US-09-105-537-30 Sequence 30, Appl
 43 48.8 1.3 36778 3 US-09-105-537-5 Sequence 5, Appl
 44 48.8 1.3 38506 3 US-09-320-878-19 Sequence 19, Appl
 45 48.8 1.3 38506 4 US-09-141-908-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
 US-08-993-228-20
 ; Sequence 20, Application US/08993228
 ; Patent No. 5976838
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Lavallie, Edward R.
 ; APPLICANT: Racie, Lisa A.
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Spaulding, Vikki
 ; APPLICANT: Agostino, Michael J.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
 ; TITLE OF INVENTION: ENCODING THEM
 ; NUMBER OF SEQUENCES: 33
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/993,228
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sprunger, Suzanne A.
 ; REGISTRATION NUMBER: 41,323
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 498-8284
 ; TELEFAX: (617) 876-5851
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3113 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-993-228-20

Query Match 10.0%; Score 370.2; DB 2; Length 3113;
 Best Local Similarity 56.4%; Pred.No. 3.4e-71;
 Matches 874; Conservative 0; Mismatches 606; Indels 69; Gaps 7;

QY 265 CCTCTGTGTGAGAGAGACCCCAACATGAAGTGTGTGTCGGCGGGGCTCAACCTGAGC 324
D 1150 GGGCGCAAGGAGATTCAGCAACACACCGAGATCTTCGGCCCCGGCGGGCGGACATGGCG 1209
QY 325 GGGTGGAGGGCCCAATGAGGACAGCGCTGGGAATGGAGTCAAGCCAGCGCCAGGCGCGAT 384
D 1210 GACAACTGGCGCAGCGAGCTGTACAAGTACAAAGTGTGTGAAGATCGAGCCCTGGGCGTG 1269
QY 385 CCCCTGACCTGCGACCGCGAGAGAGCGGAGCCCAAGAGCGGCCCAACAGTCTCCCGAG 444
D 1270 GCCCCCCACCATCGGCATCAGCAGCGGTGTGTGAGAGCGAGAGCGCGCTGACCCCTGGG- 1328
QY 445 AAGAGAGGCGCAAGGAGCTCCCTGAATGAGACGCCACCTCAGCGCGGTGTGATCTGTG 504
D 1329 --CGCCATGTTCTGGGGTTCCTGGCGCCCGCGGACGACATGTGGGCGCCCGAGCCTG 1386
QY 505 ACCAGCACCTTGACCAACAGCAAGTGTGTATCATCGACGCGCAACACCGCGGGCACGGTG 564
D 1387 ACCGTGACCGTGCAGGCCCGCGAGCTGTGAGCGGCATCGTGACAGCAGCAACAAACCTG 1446
QY 565 GTGACCAAGTTCACGCTGTGCAACCGCGACAGTGTGTGTGATCTTCAGATCCCGCGGCC 624
D 1447 CTGGCGGCATCGAGGCGCCAGCAGCACCTGTGCTGAGCTGACCGTGTGGGGCATCAAGCAG 1506
QY 625 AGCAGACGACTACCTCCCGGGAGATGTTCTTGGACAGCGAGTGAACCCAGAGGAC 684
D 1507 CTGAGGCGCGGTGTGGCGGTGAGCGCTACCTGAAAGGACAGCAGCTGTGGGCATC 1566
QY 685 CCGGGCGAGATGCGTGTGCGCGGTATCACCTGTGTGGGTGTGTCACCGCGTGCAC 744
D 1567 TGGGGCTGCAGCGCAAGCTGATCTGCACACCGCGCGTGCCTGGAGCGCGCTGGAGC 1626

RESULT 5

US-09-475-515-46

; Sequence 46, Application US/09475515A

; Patent No. 6602705

; GENERAL INFORMATION:

; APPLICANT: BARNETT, Susan

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: SRIVASTAVA, Indresh

; APPLICANT: LIAN, Ying

; APPLICANT: HARTOG, Karin

; APPLICANT: LIU, Hong

; APPLICANT: GREER, Catherine

; APPLICANT: SELBY, Mark

; APPLICANT: WALKER, Christopher

; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

; FILE REFERENCE: 1621.002

; CURRENT APPLICATION NUMBER: US/09/475,515A

; CURRENT FILING DATE: 1999-12-30

; NUMBER OF SEQ ID NOS: 90

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 46

; LENGTH: 1944

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: gpl40.mut8.modSF162.delv2

US-09-475-515-46

Query Match 1.5%; Score 55.2; DB 4; Length 1944;
Best Local Similarity 45.2%; Pred. No. 0.0077;
Matches 244; Conservative 0; Mismatches 293; Indels 3; Gaps 1;

QY 205 CCCAACGGGGCCAGAGGACACGCGATGAAGACGTCGCCGTGCGGTGTACTGCCGC 264
D 1198 CCATCCGGCGCAGATCCGTGACAGCAACATCACCGGCTGTCTGTACCCCGGAC 1257
QY 265 CCTCTGTGTGAGAGAGACCCCAACATGAAGTGTGTGTCGGCGGGGCTCAACCTGAGC 324

D 1258 GGGCGCAAGGAGATTCAGCAACACACCGAGATCTTCGGCCCCGGCGGGCGGACATGGCG 1317
QY 325 GGGTGGAGGGCCCAATGAGGACAGCGCTGGGAATGGAGTCAAGCCAGCGCCAGGCGCGAT 384
D 1318 GACAACTGGCGCAGCGAGCTGTACAAGTACAAAGTGTGTGAAGATCGAGCCCTGGGCGTG 1377
QY 385 CCCCTGACCTGCGACCGCGAAGGAGACGGGAGCCCAAGAGCGGCCCAACAGTCTCTCCCGAG 444
D 1378 GCCCCCCACCATCGCCATCAGCAGCGTGTGTGAGAGCGAGAGCGCGTGACCTGGG- 1436
QY 445 AAGAGAGGCGCAAGGAGCTCCCTGAATGAGACGCCACCTCCAGCGGGGTGTGATCTGTG 504
D 1437 --CGCCATGTTCTGTGGGCTTCCTGGGCGCGCCCGCAGCACCATGTGGGCGCCCGAGCCTG 1494
QY 505 ACCAGCACCTTGACCAACAGCAAGTGTGTATCATCGACGCGCAACACCGCGGGCACGGTG 564
D 1495 ACCGTGACCGTGCAGGCCCGCGAGCTGTGAGCGGCATGTGTGACAGCAGCAACAAACCTG 1554
QY 565 GTGACCAAGTTCACCGTGTGCAACCGCGACAGTGTGTGATCTTCAGATCCCGCGGCC 624
D 1555 CTGGCGGCATCGAGGCGCCAGCAGCACCTGTGCGAGCTGACCGTGTGGGGCATCAAGCAG 1614
QY 625 AGCAGACGACTACCTCCCGGGAGATGTTCTTGGACAGCGAGTGAACCCAGAGGAC 684
D 1615 CTGAGGCGCGCGTGTGCGCGTGTGAGCGCTACCTGAAGGACCGAGCAGCTGTGGGCATC 1674
QY 685 CCGGGCGAGATGCGTGTGCGCGGTATCACCTGTGTGGGTGTGTCACCGCGTGCAC 744
D 1675 TGGGGCTGCAGCGCAAGCTGATCTGCACACCGCGCGTGCCTGGAGCGCGAGCTGGAGC 1734

RESULT 6

US-09-475-515-45

; Sequence 45, Application US/09475515A

; Patent No. 6602705

; GENERAL INFORMATION:

; APPLICANT: BARNETT, Susan

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: SRIVASTAVA, Indresh

; APPLICANT: LIAN, Ying

; APPLICANT: HARTOG, Karin

; APPLICANT: LIU, Hong

; APPLICANT: GREER, Catherine

; APPLICANT: SELBY, Mark

; APPLICANT: WALKER, Christopher

; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

; FILE REFERENCE: 1621.002

; CURRENT APPLICATION NUMBER: US/09/475,515A

; CURRENT FILING DATE: 1999-12-30

; NUMBER OF SEQ ID NOS: 90

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 45

; LENGTH: 2025

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: gpl40.mut8.modSF162

US-09-475-515-45

Query Match 1.5%; Score 55.2; DB 4; Length 2025;
Best Local Similarity 45.2%; Pred. No. 0.0078;
Matches 244; Conservative 0; Mismatches 293; Indels 3; Gaps 1;

QY 205 CCCAACGGGGCCAGAGGACACGCGATGAAGACGTCGCCGTGCGGTGTACTGCCGC 264
D 1279 CCATCCGGCGCGCAGATCCGTGTCAGCAGCAACATCACCGGCTGTCTGTACCCCGGAC 1338
QY 265 CCTCTGTGTGAGAGAGACCCCAACATGAAGTGTGTGTCGGCGGGGCTCAACCTGAGC 324
D 1339 GGGCGCAGGAGATCAGCAACACCGAGATCTTCGGCCCCGGCGGGCGGACATGGCG 1398


```

QY 325 GGGTGGAGCCCAATGAGGACGACGCTGGGAATGGAGTCAAGCCAGCGCCAGCGCGCGAT 384
Db 1399 GACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGTGTGAAGATCGAGCCCTGGCGGTG 1458
QY 385 CCCCTGACCTGCGACCGCGAAGGAGACGCGGAGGCCCAAGAGCGCCACACGTCTCCCCGAG 444
Db 1459 GCCCCCACCATCGCCATCAGCAGCGTGTGTGACAGCGGAGAGCGCGCGTGACCCCTGGG- 1517
QY 445 AAGAAAGAGGCCCAAGGAGCTCCCTGAATGGACGCCACTCCAGCGCGGTGTGGATCTTG 504
Db 1518 --CGCATGTTCTGGGCTTCTGGGCGCGCGGAGCAGCACCATTGGGCGCCCGCAGCCTG 1575
QY 505 ACCAGACCTTGACCAACAGCAAGGTGTGTATCATCGACGCCAACAGCCGCGGACGGTG 564
Db 1576 ACCCTGACCTGACAGCGCGCGCAGCTGTGTGAGCGCATCTGTGACAGCAGACAACTTG 1635
QY 565 GTGACACAGTTACCGTCTGCAACCGGACGTCGTGTGTGTGATCTCCAGCATCCCCCGGCC 624
Db 1636 CTGGCGGCATCGAGGCCCAAGCAGCAGCCTGTGTGACGTGACCGTGTGGGCGATCAAGCAG 1695
QY 625 AGCGACAGCGACTACCTCCCGGGGAGATGTTCTTGGACAGCGACGTGAACCCAGAGAC 684
Db 1696 CTGACAGCCCGGTGCTGGCGGTGAGCGCTACCTGAAGACACAGCAGCTGTGGGCGATC 1755
QY 685 CCGGGCGCAGATGGCGTGTGCGCGGTATCATCCCTGTGGTGGGTGTGCAACCGCTGCAAC 744
Db 1756 TGGGGCTGACGGCGCAAGTATCTGCACCAACCGCGCTGCCCTGGAACGCCAGCTGGAGC 1815

```

RESULT 7

US-09-000-016-3

; Sequence 3, Application US/09000016

; Patent No. 6143541

; GENERAL INFORMATION:

; APPLICANT: Akira ARISAWA et al.

; TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ASYMMETRIC

; TITLE OF INVENTION: HYDROLASE ACTIVITY FOR 4-SUBSTITUTED 1,4-DIHYDROPYRIDINE DERIV

; TITLE OF INVENTION: ITS EXPRESSION PRODUCT

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

; STREET: 2033 K Street, N.W., #800

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Wordperfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/000,016

; FILING DATE: January 30, 1998

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Warren M. Cheek, Jr.

; REGISTRATION NUMBER: 33,367

; REFERENCE/DOCKET NUMBER:

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-721-8200

; TELEFAX: 202-721-8250

; TELEX:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2539 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: genomic DNA

; ORIGINAL SOURCE:

; ORGANISM: Streptomyces viridosporus

; STRAIN: A-914

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 338..2539

; IDENTIFICATION METHOD: E

US-09-000-016-3

Query Match

Best Local Similarity 1.5%; Score 54.8; DB 3; Length 2539;

Matches 249; Conservative 44.9%; Pred. No. 0.01;

Mismatches 302; Indels 3; Gaps 1;

QY 536 TCATCGAGCCCAACAGCCGCGGACGCGTGTGGACCAAGTTCAACCGTCTGCAACGCGCAG 595

Db 1095 TCACCGCGCGCGCGCGCGGACCAAGTGTGGGACCGGACCAACCGTCTGCAACGCGCAG 1154

QY 596 TGCTGTGATCTCCAGCATCCCGCGGACGCGGACGCGATACCTCTCCGCGGAGATGT 655

Db 1155 CGCGGGCACCGGCGCCAGTCCCAAGGGCAAGTACAAGGGCTGCGACCCGCGCGCGCA 1214

QY 656 TCCTGGACAGCGAGCTGAAACCCAGAGGACCCCGGCGCAGATGGCGTGTGGCGGTATCA 715

Db 1215 TCCTCAACGGCAAGGTCTCTCGACGACTCCGGTTTTCGGCGACGACTCCGGCATCTCGCG 1274

QY 716 CCCTGGTGGGTGTGTGCAACCGCTGCAACGCGGCGGAGCAACTGCTCTCCGAGGGG 775

Db 1275 GCATGGAGTGGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1334

QY 776 ACACCCAGTGTAGACAAAGGGGCGGCGGAGGACCCCGGCGCAGATGGCGTGTGGCGGTATCA 835

Db 1335 ACACCCAGGAGACCGCGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1394

QY 836 ACCCGTCCCGAGTCCACAGAGGAGGCGCCACAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 895

Db 1395 TCCTGTTTC---GCCATCGCGCGCGGCAACAGAGGCGCGGAGTCAATCGGTTGCGCGCGCA 1451

QY 896 GCGAGCCAGAGACAGCCATATTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 955

Db 1452 GCGCGGACGCGCGCGCTCACCGCTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1511

QY 956 CAGCCCGGACCGCGTCTCTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1015

Db 1512 CTTCCACCGCGCGCGCGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1571

QY 1016 GCAGCAGCACAGCGCCAGAGCCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1075

Db 1572 TGGACATCACGGCGCGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1631

QY 1076 CTGACCCCAATG 1089

Db 1632 CGGCGCGGCTACATG 1645

RESULT 8

US-09-514-340-3

; Sequence 3, Application US/09514340

; Patent No. 6361987

; GENERAL INFORMATION:

; APPLICANT: Akira ARISAWA et al.

; TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ASYMMETRIC

; TITLE OF INVENTION: HYDROLASE ACTIVITY FOR 4-SUBSTITUTED 1,4-DIHYDROPYRIDINE DER

; TITLE OF INVENTION: ITS EXPRESSION PRODUCT

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

; STREET: 2033 K Street, N.W., #800

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/514,340
FILING DATE: 28-Feb-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/000,016
FILING DATE: January 30, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2539 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Streptomyces viridosporus
STRAIN: A-914
FEATURE:
NAME/KEY: CDS
LOCATION: 338...2539
IDENTIFICATION METHOD: E
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-514-340-3
Query Match 1.5%; Score 54.8; DB 4; Length 2539;
Best Local Similarity 44.9%; Pred. No. 0.01;
Matches 249; Conservative 0; Mismatches 302; Indels 3; Gaps 1;
QY 536 TCATCGACGCAACAGCGGGCGCGGAGGAGTTCACCGTTCGCAACGGGACG 595
Db 1095 TCACCGCGCGCGCGCGCGGCAAGGTGGCGCAGGACCCACCGCTCGATCG 1154
QY 596 TGCTGTGATCTCCAGATCCCGCGCGGCGGAGGAGTACCTCCCGGGAGATG 655
Db 1155 CGCGGGGCAAGGCGCGCGGAGGAGGAGTACCGGCGGAGTCCGCGCGGCGA 1214
QY 656 TCCTGGACGAGCGTGAACCCAGAGGACCCGGGCGGAGTGGCGTGTGGCGGTATCA 715
Db 1215 TCCTCAACGGCAAGTCTCGACGACTCGGTTTCGGGAGACTCCGGCATCTCGCG 1274
QY 716 CCCTGGTGGGTGTGCCACCGCGTCAAGTCCCGGGGAGCAACTGTCTCTCCGAGGG 775
Db 1275 GCATGGAGTGGCGCGCGCGGCGGCGGCGGAGTCTGTCACCATGAGCTTGGCGCATGG 1334
QY 776 ACACCCGAGTGTACAGAGGGGCGGAGGAGGAGTGCACCATCGCAACGGGAAGGTCA 835
Db 1335 ACACACCGGAGGACCGCGCTGGAGGCGGCGGTGCACAGCTGTCCGCGGAGAGGGG 1394
QY 836 ACCCGTCCAGTCCACAGAGGAGGCGCACAGAGGCGGAGGTGCAGACCTTGGGCCCA 895
Db 1395 TCCTGTTCC--GCCATCGCGCGCGGCAACAGAGGGCGGAGTGCATCGGTTCCGCGGCA 1451
QY 896 GCGAGCGAGACAGCAGCAITGGCGCGCGGCGCTCTACAGAGACGCTTTCACGTACC 955
Db 1452 GCGCGGACCGCGCGCTTCAACGCTCGGCGCGGTGCAGCAGAGGACAGCTCGCGCACTTCT 1511
QY 956 CAGCCCCGACCGCGCTCTGTGGCCCCAGCTGGGAGCGAGAACGGGCGAGACCTTGACA 1015
Db 1512 CTTCAACCGGCCCCCGCTCTGGGCGAGCGGCCATCAGCGGACGTCACCGCTCCCGGG 1571
QY 1016 GCAGCAGCAGCAGGCGCAGAGCCCGAGCGGGGACCCCGAGCGGAGCAGCAGTGT 1075

Db 1572 TGGACATCAGCGCGCGCTCGCGGAGGCGCAACGACATCGCCAGGAGTCTGGTGGGAC 1631
QY 1076 CTGCACCCACCACG 1089
Db 1632 CGCGCGGCTACATG 1645
RESULT 9
US-09-000-016-1
Sequence 1, Application US/09000016
Patent No. 6143541
GENERAL INFORMATION:
APPLICANT: Akira ARISAWA et al.
TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ASYMMETRIC
TITLE OF INVENTION: HYDROLASE ACTIVITY FOR 4-SUBSTITUTED 1,4-DIHYDROPYRIDINE DERIV
TITLE OF INVENTION: ITS EXPRESSION PRODUCT
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., #800
CITY: Washington
STATE: D.C.A.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/000,016
FILING DATE: January 30, 1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2809 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Streptomyces viridosporus
STRAIN: A-914
ORIGINAL SOURCE:
ORGANISM: Streptomyces antibioticus
FEATURE:
NAME/KEY: CDS
LOCATION: 338...2539
IDENTIFICATION METHOD: E
FEATURE:
NAME/KEY: CDS
LOCATION: 2540...2809
IDENTIFICATION METHOD: P
US-09-000-016-1
Query Match 1.5%; Score 54.8; DB 3; Length 2809;
Best Local Similarity 44.9%; Pred. No. 0.011;
Matches 249; Conservative 0; Mismatches 302; Indels 3; Gaps 1;
QY 536 TCATCGACGCAACAGCGGGCGGAGGAGTTCACCGTTCGCAACGGGACG 595

Db 1095 TCACCGCCGCGCCGCGCGCGGCAACAGGTGGGCCACGCGACCCACGCTCGCTCGATCG 1154
QY 596 TGCTGTGTCATCTCCAGCATCCCGCGCGCCAGACGACGACTACCTCCCGGGGAGATGT 655
Db 1155 CGCGGGCACGGCGCCAGTCCAGGGCAAGTACAGGGCGTTCGACCCGCGCGCGCA 1214
QY 656 TCGTGACAGCGACGTGAACCCAGAGGACCCCGGCGCAGATGGCGTGTGGCCGTTATCA 715
Db 1215 TCCTCAACGGCAAGTCTCGACGACTCCGGTTTCGGCGACGACTCCGCGCTCGCGG 1274
QY 716 CCCTGTGGGTCTGTCACCCGCTGCAACGTCGCGCGGAGCAACTGTCTCTCCGAGGGG 775
Db 1275 GCATGAGTGGCG 1334
QY 776 ACACCCCATGTGTAGACAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 835
Db 1335 ACACCCGAGACCG 1394
QY 836 ACCGTCCAGTCCACAGAGGCGCACAGAGGCGCACAGGCGCGCGCGCGCGCGCGCG 895
Db 1395 TCCTGTTC---GCCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1451
QY 896 GGGAGCCAGAGACGCCCATTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 955
Db 1452 GCGCGAGCG 1511
QY 956 CAGCCCGACCG 1015
Db 1512 CTCCACCG 1571
QY 1016 GCAGAGCAGCG 1075
Db 1572 TGGACATCAGCG 1631
QY 1076 CTGACCCACCATG 1089
Db 1632 CGCGCGCTACATG 1645

RESULT 10

US-09-514-340-1
; Sequence 1, Application US/09514340
; Patent No. 6361987
; GENERAL INFORMATION:
; APPLICANT: Akira ARISAWA et al.
; TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ASYMMETRIC
; HYDROLASE ACTIVITY FOR 4-SUBSTITUTED 1,4-DIHYDROPYRIDINE DE
; ITS EXPRESSION PRODUCT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., #800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,340
; FILING DATE: 28-Feb-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/000,016
; FILING DATE: January 30, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2809 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces antibioticus
; STRAIN: <Unknown>
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 338...2539
; IDENTIFICATION METHOD: E
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2540...2809
; IDENTIFICATION METHOD: P
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-514-340-1

Query Match 1.5%; Score 54.8; DB 4; Length 2809;
Best Local Similarity 44.9%; Pred. No. 0.011;
Matches 249; Conservative 0; Mismatches 302; Indels 3; Gaps 1;

QY 536 TCATCGAGCCCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 595
Db 1095 TCACG 1154
QY 596 TGCTGTGTCATCTCCAGCATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 655
Db 1155 CG 1214
QY 656 TCCTGGAGCAGCGAGTGAACCCAGAGGACCCGCGCGCGCGCGCGCGCGCGCGCG 715
Db 1215 TCCTCAACGGCAAGTCTCGACGACTCCGGTTTCGGCGCGCGCGCGCGCGCGCGCG 1274
QY 716 CCCTGTGGGTGTGCCACCGCTGCCAACGTCGCCGCGCGCGCGCGCGCGCGCGCGCG 775
Db 1275 GCATGGAGTGGCG 1334
QY 776 ACACCCAGTGTAGACAAAGGGCGAGGGGAGGTGGCCACCATCGCCAAACGGGAAGTCA 835
Db 1335 ACACCGGAGACCGCACCCGCTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1394
QY 836 ACCCGTCCAGTCCACAGAGGCGCGCACAGAGCCACGAGGTCGAGAGCCCTGGGCCCA 895
Db 1395 TCCTGTTC---GCCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1451
QY 896 GCGAGCCAGAGACAGCCACATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 955
Db 1452 GCGCGAGCG 1511
QY 956 CAGCCCGCAGCCCGTCTCTGGCCCCCAGCGCTCGCAGCGAGAACGGGCCAGCGCTGACA 1015
Db 1512 CTCTCACCG 1571
QY 1016 GCAGCAGCACCGCCAGAGCCAGAGCCCGAGCGGGGACCCCGAGGAGCGAGGAGTGT 1075
Db 1572 TGGACATCAGCG 1631
QY 1076 CTGACCCACCATG 1089
Db 1632 CGCGCGCTACATG 1645

RESULT 11

US-09-475-515-41
; Sequence 41, Application US/09475515A
; Patent No. 6602705

```

; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 1836
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp140.mut.modSF162.delV1/V2
; US-09-475-515-41

```

```

Query Match      1.5%; Score 53.6; DB 4; Length 1836;
Best Local Similarity 45.0%; Pred. NO. 0.017;
Matches 243; Conservative 0; Mismatches 294; Indels 3; Gaps 1;

QY 205 CCCAACGGGGCCAGGAGGACACCGGATGAGACGTCGCGGTGCCGGTGTACTGCCGC 264
    |||||
Db 1090 CCCATCCCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCTGTCTGTACCCGCCAC 1149

QY 265 CCTCTGTGGAGAGGACCCACCATGAAGCTGTGTGTCGCGGGGGTCAACCTGAGC 324
    |||||
Db 1150 GGCGCAAGGAGATCAGCAACACCCAGATCTTCGCGCGGGCGGGGACATGCGC 1209

QY 325 GGGTGGAGGCCAATAGGAGACGCTGGGAATGAGTCAAGTCAAGTGTGTAAGTCCCGAT 384
    |||||
Db 1210 GACAACTGGCGCAGGAGCTGTACAAATGAGTCAAGTGTGTAAGTCCCGAT 1269

QY 385 CCCCTGACCTGCGACCGGAGAGAGACCGGAGGCCCAAGAGGCCCAACGTCCTCCCGAG 444
    |||||
Db 1270 GCCCCCAACAAAGGCCATCAGCAGCGTGTGTGAGAGGAGAGAGCGCGCTG 1326

QY 445 AAGAGAGAGGCCAAGGAGCTCCCTGAAATGAGAGCCACTCCAGCGGGTGTGATCTCTG 504
    |||||
Db 1327 GCGGCCATGTTCTGGGCTTCTGGGCGCGCGCGGAGCAACATGGGCGCGCGAGCCTG 1386

QY 505 ACCAGCACCTGACCAACAGCAAGGTGTGATCATGACGCGCAACAGCGGGGACGGTG 564
    |||||
Db 1387 ACCCTGACCGTGCAGGCGCGCAGCTGTGTGAGCGCATCTGTGAGCAGAGAACACCTG 1446

QY 565 GTGGACCAAGTTCACCGTGTGCAACCGGACGCTGTGCATCTCCAGCATCCCGCGGC 624
    |||||
Db 1447 CTGCGCGCATCGAGCGCCAGCAGCCTGTGTGAGTGTGATCTCCAGCATCCCGCGGC 684

QY 625 AGCGACAGCGATACCTCTCCCGGGAGATGTTCTTGGACAGCGAGCTGAACCCAGAGGAC 684
    |||||
Db 1507 CTGAGGCGCGGTGCTGCGCGGTGAGCGCTACCTTGAAGGACAGCAGCTGTGGGCATC 1566

QY 685 CCGGGCGCAGATGCGGTGCTGCCCGGTATCACCTCTGGTGGGTGTGCGCACCGCTGCAAC 744
    |||||
Db 1567 TGGGGTGTGAGCGGCAAGCTGATCTGCACACCGCGCTGCCCTTGGAAACGCCAGCTGGAGC 1626

```

```

RESULT 12
US-09-475-515-44
; Sequence 44, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh

```

```

; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 1836
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp140.mut7.modSF162.delV1/V2
; US-09-475-515-44

```

```

Query Match      1.5%; Score 53.6; DB 4; Length 1836;
Best Local Similarity 45.0%; Pred. NO. 0.017;
Matches 243; Conservative 0; Mismatches 294; Indels 3; Gaps 1;

QY 205 CCCAACGGGGCCAGGAGGACACCGGATGAGACGTCGCGGTGCCGGTGTACTGCCGC 264
    |||||
Db 1090 CCCATCCCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCTGTCTGTACCCGCCAC 1149

QY 265 CCTCTGTGGAGAGGACCCACCATGAAGCTGTGTGTCGCGGGGGTCAACCTGAGC 324
    |||||
Db 1150 GGCGCAAGGAGATCAGCAACACCCAGATCTTCGCGCGGGCGGGGACATGCGC 1209

QY 325 GGGTGGAGGCCAATAGGAGACGCTGGGAATGAGTCAAGTCAAGTGTGTAAGTCCCGAT 384
    |||||
Db 1210 GACAACTGGCGCAGGAGCTGTACAAATGAGTCAAGTGTGTAAGTCCCGAT 1269

QY 385 CCCCTGACCTGCGACCGGAGAGAGACCGGAGGCCCAAGAGGCCCAACGTCCTCCCGAG 444
    |||||
Db 1270 GCCCCCAACAAAGGCCATCAGCAGCGTGTGTGAGAGGAGAGAGCGCGCTG 1326

QY 445 AAGAGAGAGGCCAAGGAGCTCCCTGAAATGAGAGCCACTCCAGCGGGTGTGATCTCTG 504
    |||||
Db 1327 GCGGCCATGTTCTGGGCTTCTGGGCGCGCGCGGAGCAACATGGGCGCGCGAGCCTG 1386

QY 505 ACCAGCACCTGACCAACAGCAAGGTGTGATCATGACGCGCAACAGCGGGGACGGTG 564
    |||||
Db 1387 ACCCTGACCGTGCAGGCGCGCAGCTGTGTGAGCGCATCTGTGAGCAGAGAACACCTG 1446

QY 565 GTGGACCAAGTTCACCGTGTGCAACCGGACGCTGTGCATCTCCAGCATCCCGCGGC 624
    |||||
Db 1447 CTGCGCGCATCGAGCGCCAGCAGCCTGTGTGAGTGTGATCTCCAGCATCCCGCGGC 684

QY 625 AGCGACAGCGATACCTCTCCCGGGAGATGTTCTTGGACAGCGAGCTGAACCCAGAGGAC 684
    |||||
Db 1507 CTGAGGCGCGGTGCTGCGCGGTGAGCGCTACCTTGAAGGACAGCAGCTGTGGGCATC 1566

QY 685 CCGGGCGCAGATGCGGTGCTGCCCGGTATCACCTCTGGTGGGTGTGCGCACCGCTGCAAC 744
    |||||
Db 1567 TGGGGTGTGAGCGGCAAGCTGATCTGCACACCGCGCTGCCCTTGGAAACGCCAGCTGGAGC 1626

```

```

RESULT 13
US-09-475-515-37
; Sequence 37, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh

```

; APPLICANT: LIAN, Ying
 ; APPLICANT: HARTOG, Karin
 ; APPLICANT: LIU, Hong
 ; APPLICANT: GREER, Catherine
 ; APPLICANT: SELBY, Mark
 ; APPLICANT: WALKER, Christopher
 ; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
 ; FILE REFERENCE: 1621.002
 ; CURRENT APPLICATION NUMBER: US/09/475,515A
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 37
 ; LENGTH: 1944
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:
 ; OTHER INFORMATION: gpi40.modSF162.delV2
 US-09-475-515-37

Query Match 1.5%; Score 53.6; DB 4; Length 1944;
 Best Local Similarity 45.0%; Pred. No. 0.017;
 Matches 243; Conservative 0; Mismatches 294; Indels 3; Gaps 1;
 QY 205 CCCAACGGGGCCAGAGGACACCGCGATGAAGAACGTGCGCGTGCCTGCTGACCGCGC 264
 Db |||||
 Db 1198 CCCATCCGCGCCAGATCCGCTGCAGCAGCACATCACCGGCTGCTGACCGCGC 1257
 QY 265 CCTCTGGTGGAGAGGACCCACCATGAAGTGTGTGTCGCGGGGCGTCAACCTGAGC 324
 Db |||||
 Db 1258 GCGCGCAAGAGATCAGCAACACCAACCGAGATCTTCCGCCCGCGGGCGGACATGCG 1317
 QY 325 GGGTGGAGGCCCAATGAGCAGCGCTGGGATGAGTCAAGCAGCGCCAGCGCGCGAT 384
 Db |||||
 Db 1318 GACAACTGGCGCAGCGCTGTACAAGTACAAAGTGTGTAAGATCGAGCCCTGGGCGTG 1377
 QY 385 CCCCTGACCTCGACCGCGAGGAGACGCGCGAGCCCAAGAGCGCCACACCTCTCCCGAG 444
 Db |||||
 Db 1378 GCCCCCAAGGCCAAGCGCGGTGTGTCAGCGCGAGAA---GCGGCGCGTGAACCTG 1434
 QY 445 AAGAAGAGGCCAAGAGCTCCCTGAATGAGCGCCACCTCCAGCGGGTGTGGATCTTG 504
 Db |||||
 Db 1435 GCGGCCATGTTCTTGGGCTTCTTGGCGCGCGCGCAGCACCATGGGCGCGCGCGCGCTG 1494
 QY 505 ACCAGCACCTGACCAACAGCAAGTGTGTATCATCGACGCAACCAAGCGCGGCAACGCTG 564
 Db |||||
 Db 1495 ACCCTGACCTGAGCGCGCGCAGCTGCTGAGCGGCATCGTGACGAGCAGAAACACCTG 1554
 QY 565 GTGACCAAGTTACCGTCTGCAACGCGCAGCTGTGTGATCTCCAGATCCCGCGGCC 624
 Db |||||
 Db 1555 CTGCGGCCATCGAGGCCAGCACCTGTGCTGAGCTGACCTGCGGCGATCAAGCAG 1614
 QY 625 AGCAGACGACTACCTCCCGGGAGATGTTCTTGACAGCGAGCTGAACCCAGAGGAC 684
 Db |||||
 Db 1615 CTGACGCGCGCTGTGGCGCGTGAGCGCTACTGAAGACCAAGCAGCTGTGGGCATC 1674
 QY 685 CCGGGCGCAGATGGCGTGTGGCGGTATCACCTGTGGGCTGTGCCACCGCTGCAAC 744
 Db |||||
 Db 1675 TGGGGCTGACGCGCAAGCTGATCTGCACCAACCGCGTGCCTTGGAAACGCGCAGCTGAGC 1734

RESULT 14
 US-09-475-515-38
 ; Sequence 38, Application US/09475515A
 ; Patent No. 6602705
 ; GENERAL INFORMATION:
 ; APPLICANT: BARNETT, Susan
 ; APPLICANT: ZUR MEDEDE, Jan
 ; APPLICANT: SRIVASTAVA, Indresh
 ; APPLICANT: LIAN, Ying
 ; APPLICANT: HARTOG, Karin
 ; APPLICANT: LIU, Hong

; APPLICANT: LIU, Hong
 ; APPLICANT: GREER, Catherine
 ; APPLICANT: SELBY, Mark
 ; APPLICANT: WALKER, Christopher
 ; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
 ; FILE REFERENCE: 1621.002
 ; CURRENT APPLICATION NUMBER: US/09/475,515A
 ; NUMBER OF SEQ ID NOS: 90
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 38
 ; LENGTH: 1944
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:
 ; OTHER INFORMATION: gpi40.modSF162.delV1/V2
 US-09-475-515-38

Query Match 1.5%; Score 53.6; DB 4; Length 1944;
 Best Local Similarity 45.0%; Pred. No. 0.017;
 Matches 243; Conservative 0; Mismatches 294; Indels 3; Gaps 1;
 QY 205 CCCAACGGGGCCAGAGGACACCGCGATGAAGAACGTGCGCGTGCCTGCTGACCGCGC 264
 Db |||||
 Db 1198 CCCATCCGCGCCAGATCCGCTGCAGCAGCAACATCACCGGCTGCTGACCGCGC 1257
 QY 265 CCTCTGGTGGAGAGGACCCACCATGAAGTGTGTGTCGCGGGGCGTCAACCTGAGC 324
 Db |||||
 Db 1258 GCGCGCAAGAGATCAGCAACACCAACCGAGATCTTCCGCCCGCGGGCGGACATGCG 1317
 QY 325 GGGTGGAGGCCCAATGAGCAGCGCTGGGATGAGTCAAGCAGCGCCAGCGCGCGAT 384
 Db |||||
 Db 1318 GACAACTGGCGCAGCGCTGTACAAGTACAAAGTGTGTAAGATCGAGCCCTGGGCGTG 1377
 QY 385 CCCCTGACCTCGACCGCGAGGAGACGCGCGAGCCCAAGAGCGCCCAACCTCTCCCGAG 444
 Db |||||
 Db 1378 GCCCCCAAGGCCAAGCGCGGTGTGTCAGCGCGAGAA---GCGGCGCGTGAACCTG 1434
 QY 445 AAGAAGAGGCCAAGAGCTCCCTGAATGAGCGCCACCTCCAGCGGGTGTGGATCTTG 504
 Db |||||
 Db 1435 GCGGCCATGTTCTTGGGCTTCTTGGCGCGCGCGCAGCACCATGGGCGCGCGCGCGCTG 1494
 QY 505 ACCAGCACCTGACCAACAGCAAGTGTGTATCATCGACGCAACCAAGCGCGGCAACGCTG 564
 Db |||||
 Db 1495 ACCCTGACCTGAGCGCGCGCAGCTGCTGAGCGGCATCGTGACGAGCAGAAACACCTG 1554
 QY 565 GTGACCAAGTTACCGTCTGCAACGCGCAGCTGTGTGATCTCCAGATCCCGCGGCC 624
 Db |||||
 Db 1555 CTGCGGCCATCGAGGCCAGCACCTGTGCTGAGCTGACCTGCGGCGATCAAGCAG 1614
 QY 625 AGCAGACGACTACCTCCCGGGAGATGTTCTTGACAGCGAGCTGAACCCAGAGGAC 684
 Db |||||
 Db 1615 CTGACGCGCGCTGTGGCGCGTGAGCGCTACTGAAGACCAAGCAGCTGTGGGCATC 1674
 QY 685 CCGGGCGCAGATGGCGTGTGGCGGTATCACCTGTGGGCTGTGCCACCGCTGCAAC 744
 Db |||||
 Db 1675 TGGGGCTGACGCGCAAGCTGATCTGCACCAACCGCGTGCCTTGGAAACGCGCAGCTGAGC 1734

RESULT 15
 US-09-475-515-40
 ; Sequence 40, Application US/09475515A
 ; Patent No. 6602705
 ; GENERAL INFORMATION:
 ; APPLICANT: BARNETT, Susan
 ; APPLICANT: ZUR MEDEDE, Jan
 ; APPLICANT: SRIVASTAVA, Indresh
 ; APPLICANT: LIAN, Ying
 ; APPLICANT: HARTOG, Karin
 ; APPLICANT: LIU, Hong
 ; APPLICANT: GREER, Catherine

ug-10-019-495-26.rni

Search completed: August 24, 2004, 08:53:23
Job time : 281 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 11:10:58 ; Search time 21 Seconds
(without alignments)
1600.405 Million cell updates/sec

Title: US-10-019-495-9

Perfect score: 3473

Sequence: 1 MKNVPVPVYCRPLVEKDFM.....VLKSAERSHIIWQVSYTPE 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1446.5	41.6	828	2	US-08-993-228-21
2	146.5	4.2	985	5	PCT-US96-03916-6
3	146.5	4.2	985	5	PCT-US96-03916-66
4	143	4.1	490	3	US-09-109-841-2
5	134	3.9	406	3	US-08-582-740-68
6	134	3.9	406	3	US-09-109-879-68
7	133	3.8	382	3	US-08-582-740-70
8	133	3.8	382	3	US-09-109-879-70
9	121.5	3.5	2441	1	US-08-194-468-2
10	121.5	3.5	2441	3	US-08-961-739-2
11	121.5	3.5	2441	4	US-09-514-247A-8
12	121.5	3.5	2441	4	US-09-686-316-2
13	119.5	3.4	1231	4	US-08-714-741-41
14	119	3.4	828	4	US-09-328-352-5503
15	118.5	3.4	1038	4	US-07-757-022B-74
16	118.5	3.4	1270	4	US-07-757-022B-74
17	118.5	3.4	1311	4	US-07-757-022B-42
18	118	3.4	941	4	US-07-757-022B-14
19	118	3.4	1022	4	US-07-757-022B-84
20	118	3.4	1049	4	US-07-757-022B-58
21	118	3.4	1140	4	US-07-757-022B-104
22	118	3.4	1313	4	US-07-757-022B-142
23	118	3.4	1314	4	US-07-757-022B-50
24	118	3.4	1320	4	US-07-757-022B-46
25	118	3.4	1320	4	US-07-757-022B-60
26	118	3.4	1320	4	US-10-164-595-58
27	118	3.4	1354	4	US-07-757-022B-48

28	118	3.4	1361	4	US-07-757-022B-40
29	118	3.4	1363	4	US-07-757-022B-52
30	118	3.4	1404	4	US-07-757-022B-2
31	118	3.4	1404	4	US-07-757-022B-62
32	118	3.4	1404	4	US-10-164-595-78
33	117	3.4	1426	3	US-09-136-574A-43
34	117	3.4	8991	4	US-08-714-741-32
35	116.5	3.4	1964	2	US-08-790-912-3
36	116.5	3.4	2052	2	US-08-790-912-3
37	116	3.3	1384	3	US-08-976-255-11
38	115.5	3.3	1719	2	US-08-459-568-4
39	115.5	3.3	1719	2	US-08-399-411-4
40	115.5	3.3	1719	3	US-08-516-859A-4
41	115.5	3.3	1719	3	US-09-586-472-4
42	115.5	3.3	1719	4	US-09-528-706-4
43	114.5	3.3	325	4	US-09-510-031A-2
44	114.5	3.3	641	3	US-08-961-083-160
45	114.5	3.3	641	4	US-09-536-784-160

ALIGNMENTS

RESULT 1
US-08-993-228-21
; Sequence 21, Application US/08993228
; Patent No. 5976838
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,228
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 828 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-993-228-21

Query Match 41.6%; Score 1446.5; DB 2; Length 828;
Best Local Similarity 57.7%; Pred. No. 1.8e-126;
Matches 300; Conservative 65; Mismatches 124; Indels 31; Gaps 10;


```

; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 985 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-03916-66

```

```

Query Match          4.2%; Score 146.5; DB 5; Length 985;
Best Local Similarity 19.7%; Pred. No. 0.00018;
Matches 118; Conservative 83; Mismatches 234; Indels 165; Gaps 28;

QY 3 NVVPVYCRPLVEKDPMTKMLWCAAGVNLGWRNEDDAGNKKP--APGRPLTCDRGD 60
Db 394 SVFQETDRTLTSTPLTLPL-----TPGESE--NTLFPTAFGISTET----- 434

QY 61 GEPKSAHTSPKKKAK-----ELPMDATSSR-----VMILTSTLTTSKVVIIDANQPT 110
Db 435 --PSAAHETTQSAETVVFVTSPESEETARSQSEPPVFTQTPSTEQALTOQTIAET 492

QY 111 VVDQFTVCNAHVLCISIPAA-----SDSDYPPGEMFLDSVNPED-----PGADGVL- 158
Db 493 EALTQTPSAEQMTFTTPGAETEAAPQTPSTIP--EFTQSRSTPPETARAPSAPEVF 550

QY 159 ----AGTLVGCATRCNVPRNCSRGDTPVLDKGO--GEVATIANGVNFSQSEATE 212
Db 551 TQSSSTVEVTQTPSTVPKTLSSSTEPALFTQTQSAGTEAFTQTSAAEDTMRQTSTE 610

QY 213 A-----TEVPDPGPSEPETATLPGPLTEHVFTDPAP-----TPSSGPQPGSEN 256
Db 611 THFTQAPSTVPKATQTPSTPEVLVTSQSPSTPEVFTTLGAPEITCTPSAAPEVYTR- 669

QY 257 GPEPDSSSTRE-----PEPSGDTGAGS-----SAAPTWWLGAQNGWLYVH 298
Db 670 ----SSSTWPTAQSTPLASQNPSTSGTGHNTPEPTPVQTPHTQKLYTENKTLSPF 724

QY 299 SAVANMKCLHSIKLKDVSLSLVHK-----GRV-----L 328
Db 725 TVVSEF-----HEMTASQTPLLDVKLVVEKFSNDGEVATCTVTSKSPYRVETNWKVDL 780

QY 329 VALAD-----GTLAIFHRGEGQWDLNLYHMLDGLPHHHSIRCMVAV--YDRVWCGYKNK 381
Db 781 VDVMDEISGNSPAGVFNSE--KWQKQLYRVVTDGRTSVQMLCLSCSHSPEPCLDFTS 838

QY 382 VHVLPQKTMQIEKSFDAHPR-----ESQVQLAWIGDGVVWVSLRDLSTLRLYHAHT 433
Db 839 L-TAREXDIAPELYFTDPQATYCTITLPSGVVPRFEWSLNNVSLPEYLTATTVVSHTAG 897

QY 434 HQHLQDVDI--EPVVKMLGTGKLGFSFVRITALLVAGSRL-----WVGTGNG 479
Db 898 QSTVWKSARAGEAWIS-----GR-GNIVECTVLISDGRVTRTKERCCLTNMTIAVENG 951

```

```

RESULT 4
US-09-109-841-2
; Sequence 2, Application US/09109841
; Patent No. 6207436
; GENERAL INFORMATION:

```

```

; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Hatakeyama, Mariko
; APPLICANT: Schulein, Martin
; APPLICANT: Nielsen, Jack B.
; TITLE OF INVENTION: Endo-Beta-1, 4-Glucanases From
; FILE REFERENCE: 5195 200-US
; CURRENT APPLICATION NUMBER: US/09/109,841
; CURRENT FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: 0812/97
; EARLIER FILING DATE: 1997-07-04
; EARLIER APPLICATION NUMBER: 0846/97
; EARLIER FILING DATE: 1997-07-11
; EARLIER APPLICATION NUMBER: 60/053,506
; EARLIER FILING DATE: 1997-07-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Saccharothrix australiensis
US-09-109-841-2

```

```

Query Match          4.1%; Score 143; DB 3; Length 490;
Best Local Similarity 19.7%; Pred. No. 0.00013;
Matches 94; Conservative 40; Mismatches 153; Indels 190; Gaps 18;

QY 195 TIANGKNPSOSTEATEATEVPDPGPSEPETATLRFQPLTEHVFTDPAPTSSGPQGS 254
Db 140 TSCNGAVG-GTTEPTDEPTPEPTPEPTPEPTPEPTPEPTPEPTPEPTPEPTPEPTPEPT 191

QY 255 ENGPEPSSSTRPEPEPSCDFTGAGSAAPT-----MWLGA-----QNG 293
Db 192 EPTPEP--TPEPTPEPTPEPTPEPTPEPTPEPTPEPTPEPTPEPTPEPTPEPTPEPTPEPT 248

QY 294 WLYVHSAVANWKKLHSLKDSVLSLVHVVKGR--VIVALADGTILAIFHRGEDGQWDLN 351
Db 249 W-----VGNWNEASHAQOEVRDITSAAGAAGRTAVLVVYA-----IPGRD----- 288

QY 352 YHMLDLGPHHHSIRCMVAVYDRVWCGYKNKVHVITQPKMTQIEKSFDAHPRRESQVRQLAW 411
Db 289 -----CGQHSSEGV--STSEVAQWIDT----- 308

QY 412 IGDGV----WVSIRLDTLRLYHAHTQHLDVDIEPVVSKMLGTG-KLGFSFVRITALL 466
Db 309 VAQGI VGNPWVVL-----DPDALPMLGDCDQGDGVVGLKVAKSLT 350

QY 467 VAGSRLWVGTVGNGVVISIPLTETVVLHRLGQLLGLRANKTSPSTSGEGARPGGIIHVYGD 526
Db 351 AKGARVVIDAGHSNWL----- 367

QY 527 SDRASSFIPYCSMAQAQLCFHGHRAVKFVSVY-----GNVLATLNGS--V 572
Db 368 -----PSEAAANLNQIGFEDAVGFSINVSNVTRTTAESKWTWCQQVSLTGKKFV 416

QY 573 LDSPAEGPGAA-----PASEVSGQKLRNVLSGGEYIDFRIGDEDETEGAG 624
Db 417 IDTSRNGNGPSGSEWCNCPSGRALGERPTLVNDRSLDALLWIKLPGSDGACNGGPG 473

```

```

RESULT 5
US-08-582-740-68
; Sequence 68, Application US/08582740
; Patent No. 6037324
; GENERAL INFORMATION:
; APPLICANT: Schwender, Charles F.
; APPLICANT: Shroff, Hitesh N.
; TITLE OF INVENTION: Inhibitors of MacAM-1-Mediated
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive

```

CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/582,740
FILING DATE: 04-JAN-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-12
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 406 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-582-740-68

Query Match 3.9%; Score 134; DB 3; Length 406;
Best Local Similarity 21.7%; Pred. No. 0.00065;
Matches 91; Conservative 39; Mismatches 152; Indels 138; Gaps 22;
QY 5 PVPVYCRPL-VEKOPTMKLWCA-AGVNLGWRPNEDDAGNGVKPAGRDPLTCDREGDGE 62
Db 31 PEPVAVAGASRLQTCRLACADRGASVQ-WRGLDTSLG-AVQSDTGRSVLTVRN----- 83
QY 63 PKSAHTSPEKKAKELPEMDATSSR-----VWLSTLTTSKVIID 104
Db 84 -----ASLSAAGTRVCVCGCGRTFQHTVQLLVYAFDQLTVPAAVLV- 126
QY 105 ANQPGTVVDQFTVCNAH-----VLCISSIPAAASDSDYPPGEMFLSDVNPED---PG 153
Db 127 ---PG---DPEVACTAHKVTVPDNPALSFLLVGGQELE---GAQALGPEVQEEEEEPQG 177
QY 154 ADGVLGIT-----LVGCATRCNVPRNSCGSGDTPVLKQGQEVATIANG 199
Db 178 DEDVLFVRTERWRLPPLGTPVPVPPALYCOATWRLPGLGLSHRQAIPLVLSPTSPEPDDTTS 237
QY 200 KNPQSQTSEATEAT-----EVPDPGSRPETATLRPGPLTEHVFTDPAP---T 245
Db 238 PEPPNTTSPSPDTSPEPDDTTSQEPDDTTSQEPDDTTSQEPDDT---TSPEPDKTS 293
QY 246 PSSGPPQSGENGP-BPDSST-RPEPEPSGD-----PTG-----AGSSAAPTMWLGAQNG 293
Db 294 PEPAPQGSSTHTPRSPGSTRTRRPEISQAGPTQGEVITGSSKPADGQLPAALW----- 347
QY 294 WLYVHSVAN-----WKKCLH-----SIKLDKSVLSVHVKGRLVALA 332
Db 348 ---TSSAVLGLLLALPTVHLMKRCRHLAEDTHPPASLRLLPQVSAMAGLRGTQGVGIS 404

RESULT 6
US-09-109-879-68
Sequence 68, Application US/09109879
Patent No. 6274556
GENERAL INFORMATION:
APPLICANT: Schwender, Charles F.
APPLICANT: Shroff, Hitesh N.
TITLE OF INVENTION: INHIBITORS OF MACAM-1-MEDIATED
TITLE OF INVENTION: INTERACTIONS AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02421
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/109,879
FILING DATE: 02-JUL-1998
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/00291
FILING DATE: 03-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/582,740
FILING DATE: 04-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-12A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 406 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-109-879-68

Query Match 3.9%; Score 134; DB 3; Length 406;
Best Local Similarity 21.7%; Pred. No. 0.00065;
Matches 91; Conservative 39; Mismatches 152; Indels 138; Gaps 22;
QY 5 PVPVYCRPL-VEKOPTMKLWCA-AGVNLGWRPNEDDAGNGVKPAGRDPLTCDREGDGE 62
Db 31 PEPVAVAGASRLQTCRLACADRGASVQ-WRGLDTSLG-AVQSDTGRSVLTVRN----- 83
QY 63 PKSAHTSPEKKAKELPEMDATSSR-----VWLSTLTTSKVIID 104
Db 84 -----ASLSAAGTRVCVCGCGRTFQHTVQLLVYAFDQLTVPAAVLV- 126
QY 105 ANQPGTVVDQFTVCNAH-----VLCISSIPAAASDSDYPPGEMFLSDVNPED---PG 153
Db 127 ---PG---DPEVACTAHKVTVPDNPALSFLLVGGQELE---GAQALGPEVQEEEEEPQG 177
QY 154 ADGVLGIT-----LVGCATRCNVPRNSCGSGDTPVLKQGQEVATIANG 199
Db 178 DEDVLFVRTERWRLPPLGTPVPVPPALYCOATWRLPGLGLSHRQAIPLVLSPTSPEPDDTTS 237
QY 200 KNPQSQTSEATEAT-----EVPDPGSRPETATLRPGPLTEHVFTDPAP---T 245
Db 238 PEPPNTTSPSPDTSPEPDDTTSQEPDDTTSQEPDDTTSQEPDDT---TSPEPDKTS 293
QY 246 PSSGPPQSGENGP-BPDSST-RPEPEPSGD-----PTG-----AGSSAAPTMWLGAQNG 293
Db 294 PEPAPQGSSTHTPRSPGSTRTRRPEISQAGPTQGEVITGSSKPADGQLPAALW----- 347
QY 294 WLYVHSVAN-----WKKCLH-----SIKLDKSVLSVHVKGRLVALA 332
Db 348 ---TSSAVLGLLLALPTVHLMKRCRHLAEDTHPPASLRLLPQVSAMAGLRGTQGVGIS 404

RESULT 7
US-08-582-740-70
Sequence 70, Application US/08582740
Patent No. 6037324
GENERAL INFORMATION:

```

; APPLICANT: Schwender, Charles F.
; APPLICANT: Shroff, Hitesh N.
; TITLE OF INVENTION: Inhibitors of MacCAM-1-Mediated
; TITLE OF INVENTION: Interactions and Methods of Use Therefor
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/582,740
; FILING DATE: 04-JAN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-582-740-70

```

```

Query Match      3.8%; Score 133; DB 3; Length 382;
Best Local Similarity 22.1%; Pred. No. 0.00073;
Matches 91; Conservative 38; Mismatches 138; Indels 144; Gaps 24;

QY      5 PVPVYCRPL-VEKDPMTKLWCA-AGVNLGWRPNEDDAGNVKPAQRDPLTCDREGDGE 62
Db      31 PEPVVAVALGASRQLTCLACADRGASVQ-WRGLDTSLG-AVQSDTGRSVLTVRN----- 83
QY      63 PKSAHTSPEKKAKELPEMDATSSR-----VWILTSTLTTSKVVIIID 104
Db      84 -----ASLSAAGTRVCVCGCGGTFOHTVQLLYAFPDQLTVSPAALV- 126
QY      105 ANPGTVVDQFTVCNAH-----VLCISSIPAASDSYPPGEMFLDSVNPED---PG 153
Db      127 ---PG---DPEVACTAHKVTPVDENALSFSLVGGQELE---GAQALGPEVQEEBEPQG 177
QY      154 ADGVLAGIT-----LVGCATRCNVPRSNCSRGDTPVLDKGQGEVATIANG 199
Db      178 DEVLFRVTERWRLPPLGTVPVPALYCOATMRPLGLELSHROAIPVLH----- 225
QY      200 KVNPSQSTEEATEATEVPD-PGPSEPETATLRPGPLTEHVFTDPAP---TPSSGPGQGS 254
Db      226 --SPTSEPPEDTTSPESDTTSPESDTTSQEP-PDT-----TSPEPDKTSPEPAFQOGS 278
QY      255 ENGP-EPDSSST-RPEPEPSGD-----PTG-----AGSSAAPTMMWLGQNGWLYVHSAVA 302
Db      279 THTPRSPGSTRTRRPEISQAGPTQGEVPTGSSKPAQDQLPAALW-----TSSAVL 329
QY      303 N-----WKKCLH-----SILKDSVLSLVHVKGRVLVALA 332
Db      330 GLILLALPTVHLWKRCHLAEDDTHPPASRLLLPQVSAWAGLRTGQGVGIS 380

```

```

RESULT 8
US-09-109-879-70
; Sequence 70, Application US/09109879
; Patent No. 6274556

```

```

; GENERAL INFORMATION:
; APPLICANT: Schwender, Charles F.
; APPLICANT: Shroff, Hitesh N.
; TITLE OF INVENTION: INHIBITORS OF MacCAM-1-MEDIATED
; TITLE OF INVENTION: INTERACTIONS AND METHODS OF USE THEREFOR
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02421
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/109,879
; FILING DATE: 02-JUL-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00291
; FILING DATE: 03-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/582,740
; FILING DATE: 04-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-12A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-109-879-70

```

```

Query Match      3.8%; Score 133; DB 3; Length 382;
Best Local Similarity 22.1%; Pred. No. 0.00073;
Matches 91; Conservative 38; Mismatches 138; Indels 144; Gaps 24;

QY      5 PVPVYCRPL-VEKDPMTKLWCA-AGVNLGWRPNEDDAGNVKPAQRDPLTCDREGDGE 62
Db      31 PEPVVAVALGASRQLTCLACADRGASVQ-WRGLDTSLG-AVQSDTGRSVLTVRN----- 83
QY      63 PKSAHTSPEKKAKELPEMDATSSR-----VWILTSTLTTSKVVIIID 104
Db      84 -----ASLSAAGTRVCVCGCGGTFOHTVQLLYAFPDQLTVSPAALV- 126
QY      105 ANPGTVVDQFTVCNAH-----VLCISSIPAASDSYPPGEMFLDSVNPED---PG 153
Db      127 ---PG---DPEVACTAHKVTPVDENALSFSLVGGQELE---GAQALGPEVQEEBEPQG 177
QY      154 ADGVLAGIT-----LVGCATRCNVPRSNCSRGDTPVLDKGQGEVATIANG 199
Db      178 DEVLFRVTERWRLPPLGTVPVPALYCOATMRPLGLELSHROAIPVLH----- 225
QY      200 KVNPSQSTEEATEATEVPD-PGPSEPETATLRPGPLTEHVFTDPAP---TPSSGPGQGS 254
Db      226 --SPTSEPPEDTTSPESDTTSPESDTTSQEP-PDT-----TSPEPDKTSPEPAFQOGS 278
QY      255 ENGP-EPDSSST-RPEPEPSGD-----PTG-----AGSSAAPTMMWLGQNGWLYVHSAVA 302
Db      279 THTPRSPGSTRTRRPEISQAGPTQGEVPTGSSKPAQDQLPAALW-----TSSAVL 329
QY      303 N-----WKKCLH-----SILKDSVLSLVHVKGRVLVALA 332

```

Db 330 GLLLLALPTYHLWKRCHLAEDDTHPPASLELLPQVSAWAGLRGTGQVGLS 380

RESULT 9

US-08-194-468-2

; Sequence 2, Application US/08194468

; Patent No. 5750336

GENERAL INFORMATION:

APPLICANT: Montminy, Marc R.

TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF

TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN

TITLE OF INVENTION: RESPONSIVE GENES

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark

STREET: 444 South Flower Street, Suite 2000

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/194,468

FILING DATE: 10-FEB-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Reiter, Stephen E.

REGISTRATION NUMBER: 31,192

REFERENCE/DOCKET NUMBER: P41 9672

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619)-546-4737

TELEFAX: (619)-546-9392

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2441 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-194-468-2

Query Match 3.5%; Score 121.5; DB 1; Length 2441;

Best Local Similarity 27.2%; Pred. No. 0.17;

Matches 50; Conservative 24; Mismatches 79; Indels 31; Gaps 9;

QY 125 ISSIPAAASDSYPPGEMFLDSVNPEDPGADGVLAGITLVGCATRCN-VPRSN-C-SSRGD 182

Db 755 MASVFGMAIS---PSRM-----PQPPNMGTTHANNIMAQAQTQNFQPFSSSGA 804

QY 183 TPVLDKGQGEVATIAN-----GKVNPSQSTEEATEATEVDPD-----GPSEPETAT 228

Db 805 MSVNSVGMGQPAQAAGVSGQEPGAALPNLNLAPQASQLPCPPVTQSPHLHTPPPPAST 864

QY 229 LRPGPLTEHVFTDPAPTSSGPGQSGENGPE-PDSSSTRPEPEPSGDTGAGSSAAPTMW 287

Db 865 AAGMPSLQHP-TAPGMTP---PQPAAPTQPTSPVSSGQTPTTTPGVSVPAAQTOSTPTVQ 920

QY 288 LGAQ 291

Db 921 AAAQ 924

RESULT 10

US-08-961-739-2

; Sequence 2, Application US/08961739A

; Patent No. 6063583

GENERAL INFORMATION:

APPLICANT: Montminy, Marc R.

TITLE OF INVENTION: Methods for Treating Diabetes Mellitus

FILE REFERENCE: SALK1650-1

; CURRENT APPLICATION NUMBER: US/08/961,739A

; CURRENT FILING DATE: 1997-10-31

; EARLIER APPLICATION NUMBER: US 194,468

; EARLIER FILING DATE: 1994-02-10

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

LENGTH: 2441

TYPE: PRT

ORGANISM: Mus

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(2441)

OTHER INFORMATION: Xaa = Any Amino Acid

US-08-961-739-2

Query Match 3.5%; Score 121.5; DB 3; Length 2441;

Best Local Similarity 27.2%; Pred. No. 0.17;

Matches 50; Conservative 24; Mismatches 79; Indels 31; Gaps 9;

QY 125 ISSIPAAASDSYPPGEMFLDSVNPEDPGADGVLAGITLVGCATRCN-VPRSN-C-SSRGD 182

Db 755 MASVFGMAIS---PSRM-----PQPPNMGTTHANNIMAQAQTQNFQPFSSSGA 804

QY 183 TPVLDKGQGEVATIAN-----GKVNPSQSTEEATEATEVDPD-----GPSEPETAT 228

Db 805 MSVNSVGMGQPAQAAGVSGQEPGAALPNLNLAPQASQLPCPPVTQSPHLHTPPPPAST 864

QY 229 LRPGPLTEHVFTDPAPTSSGPGQSGENGPE-PDSSSTRPEPEPSGDTGAGSSAAPTMW 287

Db 865 AAGMPSLQHP-TAPGMTP---PQPAAPTQPTSPVSSGQTPTTTPGVSVPAAQTOSTPTVQ 920

QY 288 LGAQ 291

Db 921 AAAQ 924

RESULT 11

US-09-514-247A-8

; Sequence 8, Application US/09514247A

; Patent No. 6365361

GENERAL INFORMATION:

APPLICANT: TANABE SEIYAKU CO. LTD.

APPLICANT: TANIGUCHI, Tomoyasu

APPLICANT: MIZUKAMI, Junko

TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PP

FILE REFERENCE: TANIGUCHI=6

CURRENT APPLICATION NUMBER: US/09/514,247A

CURRENT FILING DATE: 2000-02-28

PRIOR APPLICATION NUMBER: PCT/JP98/03734

PRIOR FILING DATE: 1998-08-24

PRIOR APPLICATION NUMBER: JP231084/1997

PRIOR FILING DATE: 1997-08-27

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn version 3.0

SEQ ID NO 8

LENGTH: 2441

TYPE: PRT

ORGANISM: mouse

US-09-514-247A-8

Query Match 3.5%; Score 121.5; DB 4; Length 2441;

Best Local Similarity 27.2%; Pred. No. 0.17;

Matches 50; Conservative 24; Mismatches 79; Indels 31; Gaps 9;

QY 125 ISSIPAAASDSYPPGEMFLDSVNPEDPGADGVLAGITLVGCATRCN-VPRSN-C-SSRGD 182

Db 755 MASVFGMAIS---PSRM-----PQPPNMGTTHANNIMAQAQTQNFQPFSSSGA 804

QY 183 TPVLDKGQGEVATIAN-----GKVNPSQSTEEATEATEVDPD-----GPSEPETAT 228

Db 805 MSVNSVGMGQPAQAAGVSGQEPGAALPNLNLAPQASQLPCPPVTQSPHLHTPPPPAST 864

QY 229 LRPGPLTHVETDAFTSSGPQGSENGPE--PDSSSTRPEPEPSPGDPPTGAGSSAAPTMM 287
Db 865 AAGMPSQLHP-TAPGMTE---PQAAPTQPSTFVSQGTPPTTPGVSFAAQTSPTTVQ 920

QY 288 LGAQ 291
Db 921 AAAQ 924

```

RESULT 12
US-09-686-316-2
; Sequence 2, Application US/09686316
; Patent No. 6646115
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
; FILE REFERENCE: SALK1650-1
; CURRENT APPLICATION NUMBER: US/09/686,316
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US/08/961,739
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: US 194,468
; PRIOR FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2441
; TYPE: PRT
; ORGANISM: Mus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(2441)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-686-316-2

```

Query Match		3.5%; Score 121.5; DB 4; Length 2441;
Best Local Similarity		27.2%; Pred. No. 0.17;
Matches	50; Conservative	24; Mismatches 79; Indels 31; Gaps 9;
QY	125 ISSIPAASDDYPGEMFLSDVNPEPDGADGVLAGITLVGCATCN-VPSNC-SREGD	182
Db	: : : : - - P S R M - - - - -	:
Db	755 MASVPGMAIS- - - FSRM-	:
QY	183 TPVLDRKGQGEVIATN-----GKNVPSSOSTEEATEVEPDP-----GPSEPTETAT	228
Db	805 MSVNSVGMGGOPAAAGVSQGQEPGAALPNPLNLAPQAQSCLPCPPVTQSPULHPTTPPAST	864
QY	229 LRFGPLTEHVFTTDPAPTSSGPOPGSENKPE-PDSSSTRPEPEGPSGDTMGSSAATMW	287
Db	865 AAGMPSLQHP-TAPGMTPT---PQFAAPTQTPTPVSGQTTFPTPGSVFSAAQTSOTSTTVQ	920
QY	288 LGAQ	291
Db		:
Db	921 AAAQ	924

```

RESULT 13
US-08-714-741-41
; Sequence 41, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,741
FILING DATE: 16-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer Esq., William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2460
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1231 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-08-714-741-41

Query Match	3.4%;	Score 119.5;	DB 4;	Length 1231;	
Best Local Similarity	25.6%;	Pred. No. 0.088;			
Matches	33;	Conservative 14;	Mismatches 53;	Indels 29; Gaps 5	
QY	203	PSQSTEATEATEVPDPGPSEPEPETAIRPGPLT---	EHVFTDPAPTPSSGGPOGSENGPE	259	
		: : : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :		
Dd	652	PKPAERKAPAEKPAPAKPAEPAPKPAFATPAEAPAEQKPAPAPQAP---	APK	707	
		: : : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :		
QY	260	PDSGSSTRPEPPSGD-----	PTGAGSAAPTM-----	WLGAGNQGLYVH--	298
		: : : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :		
Dd	708	PEXPAEQPKDKTDQQQAEEEDYARRPEKPAPAPEKPAPTETPTKTGWKQENGMMFYNTD		767	
QY	299	-SAVANWKK	306		
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :		
Dd	768	GSMATGMSE	776		

```

RESULT 14
US-09-328-352-5503
; Sequence 5503, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMIN
; TITLE OF INVENTION: NUCLEIC ACID FOR DIAGNOS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5503
; LENGTH: 3892
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5503

```

Query Match 3.4%; Score 119; DB 4; Length 3892;
Best Local Similarity 20.4%; Pred. No. 0.61;
Matches 150; Conservative 76; Mismatches 286; Indels 224; Gaps 34;
QY 8 VYCRPLVEKDPDKLWCAAGVNLGWRPNEDDAGNGVKPAPGRDPTCTDREGSGEPKSAH 67

```
Db 144 VIQPTISSIEPL--LYHDAGVNPWLWAAIPLVAGIIIAAASHND-----SNDSSAPAD 195
Qy 68 TSPEKKKAKELPMDATSSRWILTSITTSKVVIIIDANOPGTVDQFTVCNAHVLCISS 127
Db 196 TTP-----PSTDGVTFVS-----DPVTSNDNVINASEASGNV-----TIITGV-----LKN 234
Qy 128 IPAASDSYPPGEMFLSDVNPDPGADGLVLAGITLVGCATRC-----NVPRNCSRG 181
Db 235 IPA-----DAANTAVTVINGVTYATVNDVKAAGTWTVSVPFSGGLVADA 277
Qy 182 DTPVLDDGQGEVATIANKNPSOSTEATEATEVP--DP-GPSEPETATLRPGPLEHV 238
Db 278 KXTIDAKVTFTDAAGNSSVNDTQYTLDTAPNAPVIDPVNGDTPDITGTAEPGSTVTVT 337
Qy 239 FTD--PAPTPSG-----POPGSENGPEPSSSTRPEPEPSGDPPTGAGSSAAATMMLGA 290
Db 338 YPDGSTKTVVAGPDGTWTVNPENGLNDGEVTAATDPAGNTSGPATAVVDAVPTV----- 393
Qy 291 QNGHLYVHSAVANWKKCLHSIKLSDSVLSL-----VHVKGRVLVALADGTLAIFHRGEGOW 347
Db 394 -----ALDDVLTNDSTPALGTVNDPTATVVNVVDVDPVAVNNGDGTW 437
Qy 348 DLSNYHMLDGLPHHSIRCM-----VVIDRWVCGYKKNKHVIQFKTMQIEKS 395
Db 438 TLADNTLPTLADGHTITVTATDAAGNVGTDGVV-----TVDTAAEPNTAGVTFT 487
Qy 396 FDAHPRRESQVRLAWIGD-----GWVWSIRLDSTLRLYHANTHQLQVDIDIEPVYSKMLG 451
Db 488 IDS--VTADNVINASEAGNVITITGVKNIPADAT-----NTATVVVINGVTYATVNDKTAG 542
Qy 452 TGKLGFSFVRITALLVAGSRLLVWGTVNGVWISLPIETV-----VLHRGQLLGL----- 500
Db 543 TWTVS-----VPGSGLVADADKTDIDAKVFTTDAAGNSSVNDTQIYTLDTAAPAA 592
Qy 501 -----RANKTSPTSGEAPRPGGIHVV-----GDD-----S 526
Db 593 PVIDPVNGTDPITGT--AEPGSTVTVTPYNGDTATVVAGPDGWSVPNPLNGDGEVEAIA 651
Qy 527 SDRASSFIPYCSMAQAQLCFHGHEDAVKFFV-----SVPGNVLATLNGSVLD 574
Db 652 TDPAGNPSLPGTATVDA-----VGPNTDGVNFTVDSVTADNVINASEASGNV--TVTGVILKN 706
Qy 575 SPAEPPGAAPASEVEGQKLRNVLSGEGV---IDFRIG-----DGED-----DET- 619
Db 707 VPADA-----ANTV-----VTVVINGQYIATVDSVTAGTWTVSVPFSGSLTADAKTI 753
Qy 620 -----EEGAGDMSQV 629
Db 754 DAKVFTTDAAGNSSSV 769

RESULT 15
US-07-757-022B-74
; Sequence 74, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1038 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-757-022B-74

Query Match 3.4%; Score 118.5; DB 4; Length 1038;
Best Local Similarity 20.2%; Pred. No. 0.083;
Matches 67; Conservative 28; Mismatches 144; Indels 93; Gaps 12;

Qy 10 CRPLVEKDPDTMKLWCAGVNLGWRPNEDDAGNGVKPAPGRDPLTCDREGDG----- 61
Db 49 CKYDKCCPDYESFCA-----EVKDNKKNRKTKKPTKPPVVDAGSLDNGDFKV 99
Qy 62 -----EPKSAHTSPKKAK-----ELPEMDATSSRVWILTS-----TLTTS 98
Db 100 TTPDTSTTQHKNVSTSPKITTAKPINRPSLPNNSDTSKETSITVKNKETTIVETKETTIN 159
Qy 99 KVVIIIDANQPGTVVDQFTVCNAHVLCISIIAASDSYPPGEMFLDSVNPEDPGADGVL 158
Db 160 KQSTDGKEKTSAK-----TQSIKTSAKDLAPTSKVLAKPT-----PKAETTT 205
Qy 159 AGITLVGCATRCNVPRSNCSRGDTPVLDDKQGEV-----ATIANGVNPSQSTBEAT 211
Db 206 KGPALT-----TPKEPTTTPKPEASTTPKEPTTIKSAPTTTPKBPAPT 252
Qy 212 EA-----TEVPDGPSEPE-----TATLRPGPLEH-----VETDPAPTSSGPOPG 253
Db 253 KSAPTTTPKEPAPTTPKEPAPTTPKEPAPTTPKBPAPTTPKBPAPTTPKBPAPT 312
Qy 254 SENGPEPDSSTRPEPEPSGDPGTGAGSSAAPT 285
Db 313 TPKEPAPTTPKEPAPTTPKBPAPT 343

Search completed: August 23, 2004, 11:15:34
Job time : 24 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 01:51:13 ; Search time 8246 Seconds
(without alignments)
13377.504 Million cell updates/sec

Title: US-10-019-495-26
Perfect score: 3694
Sequence: 1 cttgcaggttcttagtgagc.....gtcgtataaaaaaaaaa 3694

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_ptg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1478.8	40.0	2693	11	AK039130 Mus muscu
2	844.8	22.9	870	13	BQ961808 AGENCOURT
3	782.2	21.2	1146	12	BM542648 AGENCOURT
4	775.2	21.0	1201	13	EX460112 BX460112

5	772.2	20.9	1071	12	BM547756 AGENCOURT
6	746.8	20.2	868	12	BI906332 603063292
7	746.2	19.5	929	13	BQ722520 AGENCOURT
c	721.4	19.5	901	13	EX434805 BX434805
9	717	19.4	981	13	EX451957 BX451957
10	713.4	19.3	748	12	BI767677 603060824
c	683.8	18.5	690	14	CD365632 UI-H-F12
11	675.8	18.3	884	14	CD557812 AGENCOURT
12	665	18.0	721	12	BI913764 603180726
13	664.2	18.0	727	12	BI861206 603389971
14	660	17.9	862	13	BU614626 UI-M-EV0
15	642.2	17.4	1085	12	BM803227 AGENCOURT
16	637.4	17.3	827	9	AU079740 AU079740
c	634.4	17.2	690	14	CB852342 UI-CF-FN0
18	632.6	17.1	910	10	BF969625 602271988
19	628	17.0	631	14	CA422891 UI-H-FL0
c	623	16.9	978	10	BE797334 601587506
21	621.8	16.8	627	13	BQ448063 UI-H-EU1
c	618.8	16.8	622	13	EX111494 BX111494
23	618.8	16.8	1201	9	AL577398 AL577398
c	615.8	16.7	621	14	CB216188 NISC HQ02
25	613	16.6	688	10	BE907148 601499908
26	610.4	16.5	1027	13	BQ071101 AGENCOURT
27	610.4	16.5	614	12	BM666788 UI-E-CL1
c	592.4	16.0	631	13	BU633953 UI-H-FU1
c	584.2	15.8	982	12	BM449690 AGENCOURT
30	582.8	15.8	654	10	BE795312 601586616
31	582.2	15.8	587	13	BQ636226 hdb0b05.Y
c	581.6	15.7	801	14	CA319238 UI-M-FW0
33	579.6	15.7	778	14	CF536189 UI-M-G10
34	578.2	15.7	1013	13	EX451956 BX451956
35	576.6	15.6	771	14	CF724148 UI-M-GZ0
c	563.8	15.3	570	13	BQ448060 UI-H-EU1
c	561.4	15.2	571	14	CA749006 UI-H-FE1
38	558.6	15.1	875	13	BU503257 AGENCOURT
39	558.4	15.1	770	10	BE791544 601582616
40	558.4	15.1	783	14	CB518578 UI-M-GH0
41	558	15.1	748	14	CF743414 UI-M-G10
42	557.6	15.1	1091	12	BQ294455 602391568
43	555	15.0	614	10	BF952127 QV1-NN022
44	552.6	15.0	765	13	BQ572855 UI-M-PD0
45	548.6	14.9			

ALIGNMENTS

RESULT 1	AK039130	2693 bp	musculus adult male hypothalamus cDNA, RIKEN full-length protein kinase 8 interacting protein 3, full insert sequence.
LOCUS	AK039130	2693 bp	musculus adult male hypothalamus cDNA, RIKEN full-length protein kinase 8 interacting protein 3, full insert sequence.
DEFINITION	AK039130	2693 bp	musculus adult male hypothalamus cDNA, RIKEN full-length protein kinase 8 interacting protein 3, full insert sequence.
ACCESSION	AK039130	2693 bp	musculus adult male hypothalamus cDNA, RIKEN full-length protein kinase 8 interacting protein 3, full insert sequence.
VERSION	AK039130.1	2693 bp	musculus adult male hypothalamus cDNA, RIKEN full-length protein kinase 8 interacting protein 3, full insert sequence.
KEYWORDS	HTC; CAP trapper.		
SOURCE	HTC; CAP trapper.		
ORGANISM	Mus musculus		
REFERENCE	1		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		

```

3
REFERENCE
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED
11076861
4
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6
REFERENCE
AUTHORS
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imokani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahiza,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source
Location/Qualifiers
1..2693
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A230101C01"
/db_xref="MGI:2403586"
/db_xref="taxon:10090"
/clone="A230101C01"
/sex="male"
/tissue_type="hypothalamus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1764
/note="unmated protein product; mitogen-activated protein
kinase 8 interacting protein 3 (MGD|MGI:1353598,
GB|NM_013931, evidence: BLASTN, 99%, match=2586)
putative"
/codon_start=1
/protein_id="BAC30247.1"
CDS

```

```

/db_xref="GI:26333059"
/translation="STHPSPKKKAKETPEADATSSRVILTSTLTSTSKVVIIDANOP
GTIVDFQTVNCNANVLCISSIPAAASDSDDYPMGFLDSVNPDSGDGVLAGITLVGC
ATCNVPRNSRGRDTPVLIDKGGDVATTANGKVPFSQSTTEATEATEVDPDGPSS
EATTVRPGPLTEHFTDPAPSSSTQPAESNGSNGTIVQVPEPSGSLSTTISA
APTMWLCQAQNGWLYVHSANVANKKLSIKLKDVSLSLVKRGVILVALADGTIAIFH
RGSDGQWDLNSYHLMIDLGHPHSIRCMVAVNDRVMCGYKNKVHVIOPTKQWIEKSFDA
HPRESQVROLAMITGDGVVYSIRLSDTLRLYHARTHQHLQDVIDEPVSKMLGTGKLG
FSVRIITALLIAGNRLMWGTNGVVISIPITETVVLHGGQLLGIKANKTSPGSGTR
PGGIHVIYGGDSDDKAASSFIPYCSMAQQLCFHGRDRAVKFFVSVFQNVLAITNGSV
LDSPSEGGFPAAPADAEGQKLNALVLSGGBGYIDPRIGDGEDDETEECAGDVNQTK
PSLSKAERSHIIIVMQVSYTPE"
2679..2684
polyA_signal
/note="putative"
2693
polyA_site
/note="putative"
ORIGIN
Query Match 40.0%; Score 1478.8; DB 11; Length 2693;
Best Local Similarity 75.3%; Pred. No. 1.7e-234;
Matches 1970; Conservative 0; Mismatches 612; Indels 34; Gaps 9;
QY 424 AGCGCCACACGCTCTCCGAGAGAAGAGCCAGAGCTCCCTGAAATGAGCGCAC 483
DB 1 AGCACACACCCATCACCTGAGAGAAGAGCAAGCAAGCAACCCCTGAGGCGAGATGCTACC 60
QY 484 TCCAGCGGGTGTGATCTCTGACAGACACCTGACCAACAGAGTGGTGTGATCATCGAC 543
DB 61 TCCAGTGGGTATGATCTCTCACCAGCACCTGACCAACAGAGTGGTGTGATCATGAT 120
QY 544 GCCAACACAGCGGCGACGCTGTGGACCCAGTTTACCGTCTGCAACCGCCACAGTGTGTGC 603
DB 121 GCCAACACAGCGGCGACCAATTTGTGATCAGTTTCAAGTCTCAATGCCCGCTGTGT 180
QY 604 ATCTCCAGCATCCCGCGGCGCAGACAGCACTACCTCCCGGGGAGATGTTTCTGTGAC 663
DB 181 ATCTCCAGCATCTCTCGGCGCAGTGACAGTGACTATCCCTCGGGAGATGTTTCTGTAGAC 240
QY 664 AGCGACGTGAACCCAGAGACCCGGCGCAGATGGCTGCTGGCGGTATCACCCCTGTG 723
DB 241 AGTGATGTGAACCCCTGAAGATTTCAGGTGCTGATGGTGTGTGGCTGGCATCACCCCTGTG 300
QY 724 GCGTGTGCACCCCGCTGCAAGCTGCGCGGAGCAACTGCTCTCTCCGAGGGGACACCCCA 783
DB 301 GCGTGTGCTACCCGCTGCAATGTTCCAGTAGCACTGTTCTTCTACAGGAGACACCCCA 360
QY 784 GTGCTAGACAAGGGCAGGGGAGTGGCCACCATTCGCGCAACGGGAAGGTCAACCCGTCC 843
DB 361 GTACTGGACAAGGGCAGGGGAGTGGCGGACCACTGCGCAATGGGAAGGTCAACCCGTCC 420
QY 844 CAGTCCAGAGGAGGCGCACAGAGCCACGAGGTGCCAGACCTGGGCGCCAGGAGCA 903
DB 421 CAATCCAGAGAAGACCAAGACCGAGGAGTGGCCAGACCTTGGTCCAGCGAGTCA 480
QY 904 GAGACAGCCACATTCGCGGCCCGGGCTCTCTCAGAGACAGCTTCTCACTGACCCAGCCCG 963
DB 481 GNAGCAACGACAGTCCGGCCCGGGCTCTCTCAGAGAGCATGCTTTTACTGACCCAGCACCC 540
QY 964 ACCCGCTCTCTTGGCCCCCAGCTGGAGCGAGGAACGGGCGCAGAGCTTGACAGCAGCAGC 1023
DB 541 ACCCGCATCTCTCAGCACCCAGCTGCCAGTGGGTAGGATGGGTCCAGATGCGTCCAAATG 600
QY 1024 ACACGGCAGACCCAGAGCCAGCGGACCCCGAGGAGCGAGGAGGAGTGTGTCACCC 1083
DB 601 GTACAGCTCTCAGTGGAGCCCGAGTGGGAACCTCTCAACAAACAGTAGCGCTGCACCC 660
QY 1084 ACCATGTGGCTGGGAGCCCGAGAACCGCTGCTCTATGTGCACTCGGCTGTGGCCAACTGG 1143
DB 661 ACTATGTGGCTAGGAGCCCGAGAAATGGTGGCTCTATGTGCAATCAGCGGTAGCCAACTGG 720
QY 1144 AAGAAGTGGCTGCATCCATCAAGCTGAAGGATTTCTGTCTGAGCCTGCTGATGTCAAA 1203
DB 721 AAGAAGTGTGTGCTCAATCCATCAAGCTAAAAGACTCTGTGCTGAGCCTGCTGATGTCAAA 780

```


QY	1204	GCCCGTGTGCTGGTGCTCTGCGGCACGGACCCTGGGCATCTTCACCGGTGGAGAT	1263
Db	781	GGCCGAGTGCTGGTAGCTCTTCAGATGGGACCTTGGCTATCTTCATCTGTGGAGGAT	840
QY	1264	GGCAGTGGGATCTGAGCAACTATCACTTAATGGACCTGGGCCAACCGCACCATCCATC	1323
Db	841	GGCCAGTGGGACCTTGAGCAACTACCACTTAATGGACCTGGGCCAACCCACCATCCATC	900
QY	1324	CGCTCATGGCTGTGTGTACACCGCGTGTGGTGTGCTGTGTGTGTGTGTGTGTGTGTGT	1383
Db	901	CGCTCATGGCTGTGTGTGAATGACCGAGTTGGTGTGGTGTGTGTGTGTGTGTGTGTGT	960
QY	1384	ATCCAGCCCAGAACCATGCAGATAGAGAAGTCATTGTAGCGCCCAACC CGCGCGGGAGAGC	1443
Db	961	ATCCAGCCCAGACAATGCAGATTGAGAAATCAATTTGATGCCCAACCAGCGGGGAAAGC	1020
QY	1444	CAGGTGCGGACGTGGCGTGGANTGGCGAATGGGTGTTCATCCGCTGGCATCC	1503
Db	1021	CAGGTACGCTCAGCTGGCGCTGGATCGGTGATGGAGTGTGGGTCTCTATTCCGTGGATTCT	1080
QY	1504	ACCCTGAGGCTTACCATGACACAGCACACGAGCATCTACAGSAGTGGACATTGAGCCC	1563
Db	1081	ACCCTTGGCTTACCATGCTCACACCACGAGCATCTGCGAGGATGGGACATTGAGCCCC	1140
QY	1564	TACGTACGAAGATGCTAGGCACCTGGCAAGCTGGGTTTCTCTTCGTACGCATCAGGCC	1623
Db	1141	TATGTTAGCAAGATGCTAGGAACCGGCAAGCTGGGCTTCTCTTCGTGGCATCAGGCC	1200
QY	1624	CTGCTTGTGCGGGAGCCGGCTCTGGTGGGCACCGCAACGAGTGGTCATCTCCATC	1683
Db	1201	TTACTCATTTGACGGCAACCGTCTGTGGTGGGCACCTGGCAATGGGTTGTCTCATCTCCA	1260
QY	1684	CCCTTGACAGAGACTGTGGTCCTGCACCGAGCCAGCTCTCTGGGGCTCCGAGCAATAAG	1743
Db	1261	CCCTTGACTGAGACTGTGGTCCTGCATCGAGGCCAGCTCTCTAGGGCTCCGAGCAACAAG	1320
QY	1744	ACATCCCCACCTCTGGGGAGGGCGCCGTCCCGGGGGCATCATCCAGCTGATGCGCAT	1803
Db	1321	ACATCCCCAACATCTGGGGAGGGACCCGCCAGGGGGCATCATCTGTGTATGGGAC	1380
QY	1804	GACAGCAGTGACAGGGCGCCAGCAGTTTCATCCCTTACTGCTCCATGGGCCAGGCCAG	1863
Db	1381	GACAGCAGTGACAAGGGCGCCAGTAGTTTCATCCCTTACTGCTCCATGCACAGGCTCAG	1440
QY	1864	CTATGCTTCATGGCACCGCGATGCCGTGAAGTTCCTTGTCTCGGTGCCAGGGAACTGTG	1923
Db	1441	CTTTGTCTTCATGGCACCGGTGATGCTCAAATTCCTTGTCTGTGGCAGGAAATGTG	1500
QY	1924	CTGGCCACCTGATGTCAGTGTGCTGGACGCCAGCCGAGGGCCCTGGGCCAGCTGCC	1983
Db	1501	CTGGCCACTCTCAATGSCAGTGTCTAGACAGCCCATCAGAGGGCCCTGGGCTGCTGCA	1560
QY	1984	CCTGCTCGGAGTTCGAGGGCCAGAGCTCGGAACTGCTGCTGTGCTGAGCGCGGGGAG	2043
Db	1561	CCGCTGCAGATGCTGAGGGCCAGAGTTGAAGAAATGCACTGGTGTCTGAGTGGTGA	1620
QY	2044	GGCTPACATCCACTTCGCATTGGAGACGGAGAGGACGACGAGCGGAGGGCGCAGGG	2103
Db	1621	GGTTACATTGACTTCGGTATTCGGAGACGGAGAGGATGATGAACCTGAGAAATGTGCCGG	1680
QY	2104	GACATGAGCCAGGTGAAGCCCGTGTGTCAAGGCAGAGCGAGTCACATCATCTGTGTGG	2163
Db	1681	GACGTGAACCAGACAAAGCCCTCGTTGTCCAAGGCTGAGGGACGCCATCATCTGTGTGG	1740
QY	2164	CAGTGTCTTACACCCCGGAGTGAAGTGTGCTGCCCTGCTGCCCGCACCTGTACATAGGA	2223
Db	1741	CAGTGTCTTACACCCCGGAGTGAAGTGTGCTGCCCTGCTGCCCGCACCTGTACATAGGA	1799
QY	2224	CCCCGACACCTTGACCCCGCGCGCCCGCGGGGTAGCCAGCGAGCGCGCGCCCT	2283
Db	1800	CCCTTACCTGCTGCCCTCCCACTGTTCCTGTGGGCGAGCA-----GGTTCGTCCATCCC	1854

QY	2284	CTTCTAACCTCTCAACCTGCAAGCTTTTCACCTGAGTCTGGCCCCCTCCAGCGGGCAGGAGGT	2343
Db	1855	 CTTTTAACCTCTCAACTTGCAAGCTTTTGGCTTAGGTCCAGGCCCTAGCTGTGTAGAGAGGG	1914
QY	2344	GCGGGATGCGGATCACGTGGAGGAGGGGGAGGGTGTCTCCACCAGAGGGGAAGAT	2403
Db	1915	ATGAAGCAGGGTTTGGGAGTGAAGCGAGGAACAGATGCTTTTCCATTCAAGGAGGAGAGA	1974
QY	2404	GCTCTCGGGACAGTTTCCCGGCAGCTCTCTGGCCAGCTTCAGGCCACAGAT - CTTCAAGT	2462
Db	1975	GCCTCTGAGACAATCTCTCCAGCAGCCCCCTGGTCTTGACCCAATGAAGTCCAGGGGG	2034
QY	2453	CCAGGGCACATTGGGGCCCAGCGCAGGAGAAATCCGAGTGGTCTGTGCTCTACCTCTGGC	2522
Db	2035	CCCTCAGGCAAGGCAGGACAAAATGACTATGGGCAGCCCCACAGGTGGCCCTGGCTC	2094
QY	2523	CTCCTACTTCCCCAGCACCCCTGGAGGAGGAGGGGCTCCCCGCGCGGAGGCTGCTGCC	2582
Db	2095	CAC'TCCATTCCTACCATCTCTGGCAGGAGCAAGAGCCTCATACTGCCAGAGTCTTCCCCA	2154
QY	2583	CTGGGCCCACTCTGCATGCTG - CTCATGGGGCCACCCCTGCCTCTCTGGGCCCTCACTCT	2640
Db	2155	AGTCTACTTGTGTCATGCAGCCCTCAGAAAGGATCTGATTTCAAACCTTTGCTCT	2214
QY	2641	GCCTTAGGGAGCTGGGCCAGGCACTAGCTTTTGGCCAGGAGGTGGGCCCTCAGGTGCC	2700
Db	2215	CTCTGAGGGACCCAGGGCTAGSCACCATCTCTGCGCCAGAAGCAAGGTTAGGCTACCC	2274
QY	2701	AGGTGCTGCACCCCGACGGGCTCTCTGGGGCCTCCCGTCGTCAAGGCTCTATCTCG	2760
Db	2275	GGTTCCCGGGGCCACAGGAGGCTCTCT - GGGCTTCCCGCTCTCCAACTTACATTCOC	2333
QY	2761	TCTGTGCCCCACCCAGCTGTCCCTGCCAGGAGCTGGCATAAAGACACGAGGCCGCGC	2820
Db	2334	TTGTTCCCTAGCTGGTCTTACTTG - --GGNACTTGGCATGAAGCTCAGGGCCAGCC	2389
QY	2821	TCCCTGGGAGCTGCTTGAGAAACAGAGACTGTCTACCCCATCTGTGCCATGACGGCAGGC	2880
Db	2390	TCCCAAGGGCAGCTGCTT - -CAAAAGATGACAAATAACCCGTCAAACCAAGATTGTCCAGC	2447
QY	2881	TC TTGTCAGGCCCGTTCTGACCGGTGCTCCCGCCAGCTGTGCTGGGAGAGACTCACC	2940
Db	2448	CCCTCACAGACCAGTT-----CCCTCGGCTCTTCTTGACACAAGATTTCACC	2495
QY	2941	TTGGAGAGTGGGCCCTGGAGTCTCTGCTCCCAAGAGCCCCCAGGGTGGGATTCTCA	3000
Db	2496	TTGGAGAG-----CAAAGTGCCCTCCAGAGGCCCCAGGGTAGAATTCTCA	2549
QY	3001	GGCTGCCAGGCGAGGCCAGGCTCAGGAAGGG	3036
Db	2550	GGCTGCCAGGCGAGGCACAGGCTCAGGAAGGG	2585

RESULT 2

LOCUS	BQ961808	870 bp	mRNA	linear	EST 21-AUG-2002
DEFINITION	AGENCOURT_8820683 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6379940 5', mRNA sequence.				
ACCESSION	BQ961808				
VERSION	BQ961808.1 GI:22377286				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 870)				
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapbs-remail.nih.gov				
	Tissue Procurement: ATCC				
	cDNA library Preparation: Rubin Laboratory				

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM567 row: 1 column: 21 High quality sequence stop: 728. Location/Qualifiers 1. .870 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6379940" /tissue_type="epithelioid carcinoma cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 42" /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. "	
FEATURES	source
ORIGIN	
Query Match 22.9%; Score 844.8; DB 13; Length 870; Best Local Similarity 99.4%; Pred. No. 1.2e-129; Matches 857; Conservative 0; Mismatches 4; Indels 1; Gaps 1;	
QY	1142 GGAGAGTGGCTGCATCAAGCTGAGGATTCCTGTGAGCTGTGCATGCA 1201
Db	8 GGAGAGTGGCTGCATCAAGCTGAGGATTCCTGTGAGCTGTGCATGCA 67
QY	1202 AAGCCGTGTGCTGTGGCTTGGCGAGGGACCTGCCATCTCCACCGTGTGAAG 1261
Db	68 AAGCCGTGTGCTGTGGCTTGGCGAGGGACCTGCCATCTCCACCGTGTGAAG 127
QY	1262 ATGGCAGTGGATCTGACCAACTATCACTAATGGACCTGGCCACCGCACTCCA 1321
Db	128 ATGGCAGTGGATCTGACCAACTATCACTAATGGACCTGGCCACCGCACTCCA 187
QY	1322 TCCGCTGCATGCTGTTGTGACGACCGGTGTGTGGCTTACAGACAGGTGCAG 1381
Db	188 TCCGCTGCATGCTGTTGTGACGACCGGTGTGTGGCTTACAGAACAGGTGCAG 247
QY	1382 TCATCCAGCCCAAGACCATGCAGATAGAGAGTCAATTGACGCCACCGCGCGGAGA 1441
Db	248 TCATCCAGCCCAAGACCATGCAGATAGAGAGTCAATTGACGCCACCGCGCGGAGA 307
QY	1442 GCCAGTGGCGAGCTGGCGTGGATGGCGATGGCGTATGGGTGTCATCCGCTGGAT 1501
Db	308 GCCAGTGGCGAGCTGGCGTGGATGGCGATGGCGTATGGGTGTCATCCGCTGGAT 367
QY	1502 CCACCTGAGGCTTACCATGACACACGCACTTACAGACCTGAGACCTGGATGAGC 1561
Db	368 CCACCTGAGGCTTACCATGACACACGCACTTACAGACCTGAGACCTGGATGAGC 427
QY	1562 CTTAGCTGACGAAGATGCTAGGCACCTGGCAAGCTGGTTCCTTCGTACGCATCACCG 1621
Db	428 CTTAGCTGACGAAGATGCTAGGCACCTGGCAAGCTGGTTCCTTCGTACGCATCACCG 487
QY	1622 CCTGTGTTGCGCGGACCGGCTCTGGTGGGCAACCGGCAACGAGTGGTCATCTCCA 1681
Db	488 CCTGTGTTGCGCGGACCGGCTCTGGTGGGCAACCGGCAACGAGTGGTCATCTCCA 547
QY	1682 TCCCTGTGACAGACTGTTGCTCCTGACCGAGGCCAGCTCCTGGGGCTCCGAGCCAATA 1741
Db	548 TCCCTGTGACAGACTGTTGCTCCTGACCGAGGCCAGCTCCTGGGGCTCCGAGCCAATA 607
QY	1742 AGACATCCCCCACTCTGGGAGGGCGCCGCTCCCGGGGCAATCATCCAGTGTATGGCG 1801
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM567 row: 1 column: 21 High quality sequence stop: 728. Location/Qualifiers 1. .870 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6379940" /tissue_type="epithelioid carcinoma cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 42" /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. "	
FEATURES	source
ORIGIN	
Query Match 21.2%; Score 782.2; DB 12; Length 1146; Best Local Similarity 95.0%; Pred. No. 2.8e-119; Matches 852; Conservative 0; Mismatches 38; Indels 7; Gaps 4;	
QY	1704 CCTGACCGAGCGAGCTCTGGGCTCGAGCCATAGACATCCCCACCTCTGGGA 1763
Db	8 CGTGGCGGAGCGAGCTCTGGGCTCGAGCCATAGACATCCCCACCTCTGGGA 67
QY	1764 GGGCGCCCTCCCGGGGGCATCATCCAGTGTATGGCGATGACAGTGCAGGCGCG 1823
Db	68 GGGCGCCCTCCCGGGGGCATCATCCAGTGTATGGCGATGACAGTGCAGGCGCG 127
QY	1824 CAGCAGTTTCATCCCTACTGTCTCCATGGCCCGAGCCAGCTATGTTTCATGGGCA 1983


```
Db      717  GGCTCTCCCTGCTCAAGCCTCTATCTCTGCTGTCTCCCAACCCCAAGCTGTTCCCTGCCCCAG 776
Qy      2792  GGAGCTGGCATAAAGACAGAGGCCCGGCTCCCTGGGGGAGCTGCTTTGAGAACAGAGACT 2851
Db      777  GGAGCTGGCATAAAGACAGAGGCCCGGCTCYCTGGGGGAGCTKCTTTGAGAACARAGACT 836
Qy      2852  GCTACCCCACTCTGCCCATGAGCAGGCTCTTGGCAGCCCCCGTTCTGACCCCGTGTCCCC 2911
Db      837  GYTACCCCATCTGSCCATGAGAGGCTGTTTSCMACCCCGTCTGTGACGGGK---CCC 893
Qy      2912  CCAGGCTCTGCTGGGAGAGACTCACCTTTGGAG 2946
Db      894  CCARGYKCTG-CTGGGAGAAAAATCACTTTGSAG 927

RESULT 5
BM547756      1071 bp      mRNA      linear      EST 20-FEB-2002
LOCUS      AGENCOURT_6507064 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5727893
DEFINITION      5', mRNA sequence.
ACCESSION      BM547756
VERSION
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1071)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12722 row: 1 column: 06
High quality sequence stop: 676.

FEATURES
source
location/Qualifiers
1..1071
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5727893"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
(destroyed); Site_2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones, and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."

ORIGIN

Query Match      20.9%; Score 772.2; DB 12; Length 1071;
Best Local Similarity 97.6%; Pred. No. 1.2e-117;
Matches 805; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

Qy      666  CGACGTGAACCCAGAGGACCCGGGCGCAGATGCGTGTGGCGGTATCACCCCTGGTGGG 725
Db      1  CGACGTGAACCCAGAGGACCCGGGCGCAGATGCGTGTGGCGGTATCACCCCTGGTGGG 60
Qy      726  CTGTGCGCACCCGCTGCAACGTGCGCGGAGCACTCTCTCCCGAGGGGACACCCCACT 785
Db      61  CTGTGCGCACCCGCTGCAACGTGCGCGGAGCACTCTCTCCCGAGGGGACACCCCACT 120

786  GTAGACAAGGGGAGGGGAGGTGGCCACCATCTGCCAACAGGGAAGTCAACCGTCCCA 845
121  GCTTAGACAAGGGGAGGGGAGGTGGCCACCATCGCAACGGGAAGGTCAACCGTCCCA 180
846  GTCCACAGAGAGGCCCAACAGAGGCCACGAGAGTGGCCAGACCTTGGGGCCACGAGCCAGA 905
181  GTCCACAGAGAGGGCCCAACAGAGGCCACGAGAGTGGCCAGACCTTGGGGCCACGAGCCAGA 240
906  GACAGCCACATTTGGGCCCCGGGCTCTCACAGAGCAGTCTTCACTGACCCAGCCCGAC 965
241  GACAGCCACATTTGGGCCCCGGGCTCTCACAGAGCAGTCTTCACTGACCCAGCCCGAC 300
966  CCGTCTCTTGGCCCCCAGAGCTTGGCAGCGAGAACGGGCCAGAGCTGACAGCAGCAGC 1025
301  CCGTCTCTTGGCCCCCAGAGCTTGGCAGCGAGAACGGGCCAGAGCTGACAGCAGCAGC 360
1026  ACGGCCAGAGCCAGAGCCCGAGCGGGAGCCCAACGGGAGCAGGAGCAGTGTGACCCAC 1085
361  ACGGCCAGAGCCAGAGCCCGAGCGGGAGCCCAACGGGAGCAGGAGCAGTGTGACCCAC 420
1086  CATGTGGCTGGAGCCCGAGAGCGGCTGCTATGTGCACTCGCTGTGGGCCAACTGGAA 1145
421  CATGTGGCTGGAGCCCGAGAGCGGCTGCTATGTGCACTCGCTGTGGGCCAACTGGAA 480
1146  GAAGTGGCTGCACTCCCATCAAGCTGAAGGATTTCTGTGTGAGCCTGTGGTGCATCGA 1205
481  GAAGTGGCTGCACTCCCATCAAGCTGAAGGATTTCTGTGTGAGCCTGTGGTGCATCGA 540
1206  CCGTGTGTGTGGCTCTTGGCGGAGCGGACCCCTGGCCATCTTCCACCGTGTGAGATGG 1265
541  CCGTGTGTGTGGCTCTTGGCGGAGCGGACCCCTGGCCATCTTCCACCGTGTGAGATGG 600
1266  CCAGTGGGATCTGAGCACTATCACTAATGAGCTGTGGGCCACCGGACCACTCCATCCG 1325
601  CCAGTGGGATCTGAGCACTATCACTAATGAGCTGTGGGCCACCGGACCACTCCATCCG 660
1326  CTGCATGCTGTTGTGTACGACCGCGTGTGGTGTGGCTTACAAGAACAAAGGTGACGTAT 1385
661  CTGCATGCTGTTGTGTACGACCGCGTGTGGTGTGGCTTACCAGAACAAAGGTGACGTAT 720
1386  CCAGCCCAAGACCATGACATAGAGAGTCAATTGACG-CCACCCGCGGGGAGAGCC 1444
721  CCAGCCCAAGACCATGACATAGAGAGTCAATTGACG-CCACCCGCGGGGAGAGCC 780
1445  AGGTGCGCAGCTGCGGTGG-ATCGCGCATGCGGTATGGGTGCC 1488
781  CAGTGGCGCAACTTGGGTTGGAATCGGCAAGGGCTTATGGTTGCC 825

RESULT 6
BI906332
LOCUS      603063292F1 NIH_MGC_118 Homo sapiens CDNA clone IMAGE:5212463 5',
DEFINITION      mRNA sequence.
ACCESSION      BI906332
VERSION      BI906332.1 GI:16169009
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 868)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
cDNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
```

http://image.llnl.gov
Plate: LLAM1533 row: g column: 24
High quality sequence stop: 812.
Location/Qualifiers
1..868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5212463"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN
Query Match 20.2%; Score 746.8; DB 12; Length 868;
Best Local Similarity 97.2%; Pred. No. 1.9e-113;
Matches 824; Conservative 0; Mismatches 17; Indels 7; Gaps 6;

QY 1096 GGAGCCACAGACGGCTCTATGTGACCTGGCTGGCCAACTGGAGNAGTGCCTG 1155
Db 1 GGAGCCACAGACGGCTCTATGTGACCTGGCTGGCCAACTGGAGNAGTGCCTG 60

QY 1156 CACTCCA-TCAAGTGAAGGATCTGTGCTGAGCCTGGTGCATGTCAAAGGCCGTGCT 1214
Db 61 CACTCCAGTCAAGTGAAGGATCTGTGCTGAGCCTGGTGCATGTCAAAGGCCGTGCT 120

QY 1215 GGTGCTCTGGCGACGGACCTGGCCCATCTTCCACCGTGTGAAGTGCAGTGGGA 1274
Db 121 GGTGCTCTGGCGACGGACCTGGCCCATCTTCCACCGTGTGAAGTGCAGTGGGA 180

QY 1275 TCTGAGCAACTATCACTTAATGAGCTGGCCACCCGACCACTCCATCCGCTGATGGC 1334
Db 181 TCTGAGCAACTATCACTTAATGAGCTGGCCACCCGACCACTCCATCCGCTGATGGC 240

QY 1335 TGTGTGTACACCGCTGTGTGTGTGCTACAAGAACAGGTGCACGTCAATCCAGCCAA 1394
Db 241 TGTGTGTACACCGCTGTGTGTGTGCTACAAGAACAGGTGCACGTCAATCCAGCCAA 300

QY 1395 GACCATGAGATAGAGATCATTTGACGCCACCCGCGCGGAGAGCCAGGTGCGCA 1454
Db 301 GACCATGAGATAGAGATCATTTGACGCCACCCGCGCGGAGAGCCAGGTGCGCA 360

QY 1455 GCTGGCGTGGATCGGATGCGTATGCGTGTGCTCATCGCTGACCTCCACCTGAGGCT 1514
Db 361 GCTGGCGTGGATCGGATGCGTATGCGTGTGCTCATCGCTGACCTCCACCTGAGGCT 420

QY 1515 CTACCATGCAACACGACCACTATACAGGAGTGGACATTCAGGCTTACGTCAGCAA 1574
Db 421 CTACCATGCAACACGACCACTATACAGGAGTGGACATTCAGGCTTACGTCAGCAA 480

QY 1575 GATGCTAGGCACTGGCAAGCTGGTTTCTTCTGTAAGCATACGCGCTGTGTCGC 1634
Db 481 GATGCTAGGCACTGGCAAGCTGGTTTCTTCTGTAAGCATACGCGCTGTGTCGC 539

QY 1635 GGGCAGCCGGCTCTGGTGGGACCGGCAAGGATGTGTCATCTCCATCCCCCTGACAGA 1694
Db 540 GGGCAGCCGGCTCTGGTGGGACCGGCAAGGATGTGTCATCTCCATCCCCCTGACAGA 599

QY 1695 GACTGTGGTCTGACCGAGCCAGCTCTCTGGGCTCCGAGCCAAATAGACATCCCCAC 1754
Db 600 GACTGTGGTCTGACCGAGCCAGCTCTCTGGGCTCCGAGCCAAATAGACATCCCCAC 659

QY 1755 CTCTGGGAGGGGCGCCGCTCCCGGGGCATCATCCAGTGTATGGCGATGACAGCAGTGA 1814

Db 1 AGTCATTGACGCCACCCCGCGGAGAGCAGGTGCGGCTGGCTGGATCGGCG 60

1660 CTCTGGGAGGGGCGCCGCTCCCGGGGCATCATCCAGTGTATGGCGATGACAGCAGTGA 719

QY 1815 CAGGGCGGCGCAGCAGCTTCCATCCCTTCTGCTCCATGGCCCGCAGCTGCTTCCA 1874

Db 720 CAGGGCGGCGCAGCAGCTTCCATCCCTTCTGCTCCATGGCCCGCAGCTGCTTCCA 775

QY 1875 TGGGACCCCGCATGCGGTGAGTTCTTGTCTCGGTGCGCAGGAAAGTGTGGCCACCC 1933

Db 776 TGGGACCCCGCATGCGGTGAGTTCTTGTCTCGGTGCGCAGGAAAGTGTGGCCACCC 835

QY 1934 TGAATGGC 1941

Db 836 TGAATGGC 843

RESULT 7
LOCUS BQ722520
DEFINITION AGENCOURT_8305079_lupski_sympathetic_trunk_Homo_sapiens_cDNA_clone
IMAGE:6193163_5', mRNA sequence.
ACCESSION BQ722520
VERSION BQ722520.1 GI:21861417
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rcgaps@remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13596 Row: f Column: 12
High quality sequence stop: 592.
Location/Qualifiers
1..929
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6193163"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski_sympathetic_trunk"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1: NotI; Site 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors:
5'-TCGACCCAGCCGTCG-3' and
5'-GACTAGTCTTAGTCGCGAGCGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."

ORIGIN
Query Match 20.2%; Score 746.2; DB 13; Length 929;
Best Local Similarity 90.4%; Pred. No. 2.4e-113;
Matches 831; Conservative 0; Mismatches 83; Indels 5; Gaps 3;

QY 1412 AGTCATTGACGCCACCCCGCGGAGAGCAGGTGCGGCTGGCTGGATCGGCG 1471

Db 1 AGTCATTGACGCCACCCCGCGGAGAGCAGGTGCGGCTGGCTGGATCGGCG 60

ACCESSION BI767677
 VERSION BI767677.1 GI:15759255
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 748)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L14M11526 row: p column: 05
 High quality sequence stop: 746.
 Location/Qualifiers
 1..748
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5209972"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 122"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NO1; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 19.3%; Score 713.4; DB 12; Length 748;
 Best Local Similarity 99.5%; Pred. No. 6.3e-108;
 Matches 747; Conservative 0; Mismatches 1; Indels 3; Gaps 3;
 QY 1125 CTCGGCTGGCCAACTGAAGAGTGGCTGCACCTCCATCAAGCTGAAGGATTTGTGCT 1184
 DB 1 CTCGGCTGGCCAACTGAAGAGTGGCTGCACCTCCATCAAGCTGAAGGATTTGTGCT 60
 QY 1185 GAGCCTGGTGATGTCAAAAGGCGGTGTGTGGCTCTGGCGGACGGACCCCTGGCCAT 1244
 DB 61 GAGCCTGGTGATGTCAAAAGGCGGTGTGTGGCTCTGGCGGACGGACCCCTGGCCAT 120
 QY 1245 CTTCCACCGTGTGAAGTGGCAGTGGATCTGAGCAACTATCACTTAATGACCTGGG 1304
 DB 121 CTTCCACCGTGTGAAGTGGCAGTGGATCTGAGCAACTATCACTTAATGACCTGGG 180
 QY 1305 CCACCGGCACCACTCCATCCGCTGCATGGCTGTGTGACGCGCGTGTGTGGCTA 1364
 DB 181 CCACCGGCACCACTCCATCCGCTGCATGGCTGTGTGACGCGCGTGTGTGGCTA 240
 QY 1365 CAAGACAGGTGCAGTCTATCCAGCCAGACCACTGAGATGAGAGTCAATTTGACGC 1424
 DB 241 CAAGACAGGTGCAGTCTATCCAGCCAGACCACTGAGATGAGAGTCAATTTGACGC 300
 QY 1425 CCACCGCGCGGAGAGCCAGTGGCGGAGTGGCGTGGATCGGCGATGGGTATGGGT 1484
 DB 301 CCACCGCGCGGAGAGCCAGTGGCGGAGTGGCGTGGATCGGCGATGGGTATGGGT 360
 QY 1485 GTCCATCCGCTGGACTCCACCTGAGGTCTACCATGACACACGACCATCTACA 1544
 DB 361 GTCCATCCGCTGGACTCCACCTGAGGCTTACCATGACACACGACCATCTACA 420

QY 1545 GGACGTGGACATTGAGCCCTACGTCAGCAAGATGCTAGGCACCTGGCAAGCTGGGTTTCTC 1604
 DB 421 GGACGTGGACATTGAGCCCTACGTCAGCAAGATGCTAGGCACCTGGCAAGCTGGGTTTCTC 480
 QY 1605 CTTCTGACGATCAGCGCCCTGCTGTGTGCGGGCAGCCGCTCTGGGTGGGCACCGGCAA 1664
 DB 481 CTTCTGACGATCAGCGCCCTGCTGTGTGCGGGCAGCCGCTCTGGGTGGGCACCGGCAA 540
 QY 1665 CGGAGTGGTATCTCCATCCCTGACAGAGACTGTGCTGTGACCGAGGCCAGCTCCT 1724
 DB 541 CGGAGTGGTATCTCCATCCCTGACAGAGACTGTGCTGTGACCGAGGCCAGCTCCT 600
 QY 1725 GGGCTCCGAGCCAAATAAGACATCCCCACCTCTGGGGAGGGCCCGCTCCCGGGGGCAT 1784
 DB 601 GGGCTCCGAGCCAAATAAGACATCCCCCA-CTCTGGGGAGGGCCCGCTCCCGGGGGCAT 659
 QY 1785 CATCCAGTGTATGGCGATGACAGAGTGTGACAGGGCGGCAGAGCTTCATCCCTACTG 1844
 DB 660 CATCCAGTGTATGGCGATGACAGAGTGTGACAGGGCGGC-CAGCAGCTTCATCCCTACTG 718
 QY 1845 CTCCATGGCCAGGCCCGCAGTATGCTTCCAT 1875
 DB 719 CTCCATGGCCAAAGG-CCAGCTATGCTTCCAT 748
 RESULT 11
 CD365632/c
 LOCUS
 DEFINITION UI-H-Ft2-bjk-a-24-0-UI.s1 NCI CGAP_Ft2 Homo sapiens cDNA clone
 UI-H-Ft2-bjk-a-24-0-UI 3', mRNA sequence.
 ACCESSION CD365632
 VERSION CD365632.1 GI:31149722
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 690)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Gary W. Hunninghake, U of I
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at <http://genome.uiowa.edu/distribution/cgap.html>
 Seq primer: M13 FORWARD
 POLYA=yes.
 Location/Qualifiers
 1..690
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-Ft2-bjk-a-24-0-UI"
 /tissue_type="Adult"
 /dev_stage="Aveolar Macrophage"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP_Ft2"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified Polylinker; Site 1: EcoR 1; Site 2: Not I; NCI CGAP_Ft2 is a subcloned cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was subcloned according to Ronaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.
 TAG_TISSUE=Human Lung Aveolar Macrophage
 TAG_LIB=UI-H-Ft2


```
ORIGIN
Query Match      18.5%; Score 683.8; DB 14; Length 690;
Best Local Similarity 99.6%; Pred. No. 4.9e-103;
Matches 685; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
TAG_SEQ=GGCCATGCCG"

QY 3007 CAGGCGAGCCCGAGGCTCAGCAAGAGGGAGGCGCCCTGGGCTCTCCGGGATCAGTCT 3066
DB 690 CAGGCGAGCCCGAGGCTCAGCAAGAGGGAGGCGCCCTGGGCTCTCCGGGATCAGTCT 631
QY 3067 AGGACACAGGCTCAGGCTCAGGTTGATGGGGATGATGTCTCCCGGGGCTGCCTCGT 3126
DB 630 AGGACACAGGCTCAGGCTCAGGTTGATGGGGATGATGTCTCCCGGGGCTGCCTCGT 571
QY 3127 CACGGGGTCCACGAGGCCAGCTCCACACAGCTACTAAGTGTCTAGGTTGGCCGCT 3186
DB 570 CACGGGGTCCACGAGGCCAGCTCCACACAGCTACTAAGTGTCTAGGTTGGCCGCT 511
QY 3187 GTGGCTCTCTCCAGGAGCAACAGAGAGGCGCCACCAAGCAGAGGCGCTGGGGCTGAGGA 3246
DB 510 GTGGCTCTCTCCAGGAGCAACAGAGAGGCGCCACCAAGCAGAGGCGCTGGGGCTGAGGA 451
QY 3247 TGGAGCCGCCCGAGCGACTCCAAAGCCCGCAGAGGGCAGCGCCACCTTGGACTGTCT 3306
DB 450 TGGAGCCGCCCGAGCGACTCCAAAGCCCGCAGAGGGCAGCGCCACCTTGGACTGTCT 391
QY 3307 CCTGCCAGCTGGGCTCTCTGGCTATCTTACCTTCCAGGCCCACTGCACCTCTCTC 3366
DB 390 CCTGCCAGCTGGGCTCTCTGGCTATCTTACCTTCCAGGCCCACTGCACCTCTCTC 331
QY 3367 TGGAGGCCCTTATAGGAGCGCCAGCGCCCGCCAGCCACCCACCAAGCAGAGGAGCAG 3426
DB 330 TGGAGGCCCTTATAGGAGCGCCAGCGCCCGCCAGCCACCCACCAAGCAGAGGAGCAG 271
QY 3427 ATCTTGGGAGCTGCCCAAGCCCGCTGCCCAAGCCCGCTGCCCAAGCCCGCTGCCGTC 3486
DB 270 ATCTTGGGAGCTGCCCAAGCCCGCTGCCCAAGCCCGCTGCCCAAGCCCGCTGCCGTC 211
QY 3487 CGGCTTCTCCCAACACCTCCACCTCCACCTCCATGTATGTATGTCTCCCTCTGTGT 3546
DB 210 CGGCTTCTCCCAACACCTCCACCTCCATGTATGTATGTCTCCCTCTGTGT 151
QY 3547 CCCCAGGATCTGAAGTGAATCCGGGCTGAGCAAGTGGGGGGCTGGGGAGGGGTGACG 3606
DB 150 CCCCAGGATCTGAAGTGAATCCGGGCTGAGCAAGTGGGGGGCTGGGGAGGGGTGACG 91
QY 3607 ATTCTCTCAGGCTTGGCCCTCAAGCAAAACCCACATATCTGTATGTATTAATAAT 3666
DB 90 ATTCTCTCAGGCTTGGCCCTCAAGCAAAACCCACATATCTGTATGTATTAATAAT 31
QY 3667 GTCTTAACGTCGTAATAAAAAAAAAAAAA 3694
DB 30 GTCTTAACGTCGTAATAAAAAAAAAAAAA 3

RESULT 12
CD557812
LOCUS
DEFINITION
AGENCOURT_14413288 NIH_MGC_180 Homo sapiens cdna clone
IMAGE:30386864 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 884)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Daniela S. Gerhard, Ph.D.
```

Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM456 row: d column: 09
High quality sequence stop: 550.
Location/Qualifiers
1. 884
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30386864"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_180"
/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site.1: NotI;
Site.2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

FEATURES
source

ORIGIN

Query Match 18.3%; Score 675.8; DB 14; Length 884;
Best Local Similarity 97.9%; Pred. No. 1.1e-101;
Matches 706; Conservative 0; Mismatches 12; Indels 3; Gaps 2;
QY 1825 AGCAGCTTCATCCCTACTGTCTCATGCCCCAGGCCAGCTATGCTTCCATGGGACCGC 1884
DB 1 AGCAGCTTCATCCCTACTGTCTCATGCCCCAGGCCAGCTATGCTTCCATGGGACCGC 60
QY 1885 GATCCCGTGAAGTTCTTTGTCTCGGTGCCAGGGAACGTGTGGCCACCCCTGAATGGCAGT 1944
DB 61 GATCCCGTGAAGTTCTTTGTCTCGGTGCCAGGGAACGTGTGGCCACCCCTGAATGGCAGT 120
QY 1945 GTGCTGACAGCCCGAGGCGCCCTGGGCGAGCTGCCCTGCTCGAGGTGAGGGC 2004
DB 121 GTGCTGACAGCCCGAGGCGCCCTGGGCGAGCTGCCCTGCTCGAGGTGAGGGC 180
QY 2005 CAGAAGCTGCGGAACGTGTGTGTCTGAGCGGGGAGGAGGCTACATCGACTTCCGCAAT 2064
DB 181 CAGAAGCTGCGGAACGTGTGTGTCTGAGCGGGGAGGAGGCTACATCGACTTCCGCAAT 240
QY 2065 GGAGCGGAGGAGCAGCAGCAGAGCGAGGAGGGCGCAGGGACATGAGCCAGGTGAAGCCC 2124
DB 241 GGAGCGGAGGAGCAGCAGCAGAGCGAGGAGGGCGCAGGGGACATGAGCCAGGTGAAGCCC 300
QY 2125 GTGCTGTCCAAAGCAGAGCGCAGTCATCATCTGTGTGGGAGGTGCTTACACCCCGAG 2184
DB 301 GTGCTGTCCAAAGCAGAGCGCAGTCATCATCTGTGTGGGAGGTGCTTACACCCCGAG 360
QY 2185 TGAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2244
DB 361 TGAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 2245 CCGGCGCGCGGGGTAGCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2304
DB 421 CCGGCGCGCGGGGTAGCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 2305 GCTTTTCCACCTGAGTCTGGGCGCTTCCAGCGGCGAGGAGTGGCGGGATGCGGATCAGTGG 2364
DB 481 GCTTTTCCACCTGAGTCTGGGCGCTTCCAGCGGCGAGGAGTGGCGGGATGCGGATCAGTGG 540
QY 2365 GAGGAGGAGGAGGAGGAGTGTCTTCCACCCGAGGAGGAGAGTGTCTCGGGAGAGTTTCCCGG 2424
DB 541 GAGGAGGAGGAGGAGGAGTGTCTTCCACCCGAGGAGGAGAGTGTCTTCCGGAGAGTTTCCCGG 600

```

QY 2425 GCAGCTCTGGCCAGCTTCCAGCCACAGCTCCTCAAGTCCAGGCGACCTTGGGCGCCAGCG 2484
Db 601 GCAGCTCTGGCCAGCTTCCAGCCACAGCTCCTCAAGTCCAGGCGACCTTGGGCGCCAGCG 660
QY 2485 CAGSCAGAAT-CCGAGGTGGTCTGGCTTACCC- -TGGGGCTCTACTCTCCCGCAGCACCC 2541
Db 661 CAGCAGAAATCCCAAGGTGGTCTGGGCGCCCTCCCGCTGGGGCTCTACTCTCCCGCAGCACCC 720
QY 2542 C 2542
Db 721 C 721

RESULT 13
BI913764
LOCUS 603180726F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5244630 5',
DEFINITION mRNA sequence.
ACCESSION BI913764
VERSION BI913764.1 GI:16178127
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 721)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11617 row: d column: 07
High quality sequence stop: 716.
FEATURES
Location/Qualifiers
1..721
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="IMAGE:5244630"
/lab_host="DH10B"
/clone_lib="NIH_MGC_121"
/note="organ: brain; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match 18.0%; Score 665; DB 12; Length 721;
Best Local Similarity 98.8%; Pred. No. 6.4e-100;
Matches 712; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
QY 1787 TCCAGTGTATGGCGATGACAGCAGTGACAGGCGCGCCAGCAGCTTCATCCCTACTGCT 1846
Db 1 TCCAGTGTATGGCGATGACAGCAGTGACAGGCGCGCCAGCAGCTTCATCCCTACTGCT 60
QY 1847 CCATGCCCGAGCCCGCAGCTATCTCTCCATGGCACCGCATGCCGTAAGTCTTTGTCT 1906
Db 61 CCATGCCCGAGCCCGCAGCTATCTCTCCATGGCACCGCATGCCGTAAGTCTTTGTCT 120
QY 1907 CGGTGTCAGGGAACGTGCTGGCCACCCTGAATGGCAGTGTGCTGGACAGCCCGCAGG 1966

```

```

Db 121 CGGTGCCAGGGAAGTGTGCTGGCCACCTGAATGGCAGTGTGCTGGACAGCCCGCAGG 180
QY 1967 GCGCTGGGCGAGCTGCCCCCTGCTCGAGAGTCGAGGCGCAGAACTCGGAACTGCTGG 2026
Db 181 GCGCTGGGCGAGCTGCCCCCTGCTCGAGAGTCGAGGCGCAGAACTCGGAACTGCTGG 240
QY 2027 TGCTGAGCGGCGGAGGAGGCTACATCGACTTCCGCAATTGGAGACGAGAGACACAGGA 2086
Db 241 TGCTGAGCGGCGGAGGAGGCTACATCGACTTCCGCAATTGGAGACGAGAGACACAGGA 300
QY 2087 CGGAGGAGGCGCAGGCGGACATGAGCCAGGTGAAGCCGCTGCTTCCAAAGCAGAGCGCA 360
Db 2147 GTCAATCATCTGTGTGAGAGTGTCTTACACCCCGGAGTGAAGCTGCTGCCCTGCCCTGGC 2206
Db 361 GTCAATCATCTGTGTGAGAGTGTCTTACACCCCGGAGTGAAGCTGCTGCCCTGCCCTGGC 420
QY 2207 CCGACCTGTATACAGGACCCCGGACCACTTACACCCCGGCGCGCGGGGTAGCCAGC 2266
Db 421 CCGACCTGTATACAGGACCCCGGAGCACTTACACCCCGGCGCGGGGTAGCCAGC 480
QY 2267 CAGCGCGCGCGCGCGCTTCTTAACTCTCAACCTGACAGCTTTCACCTGAGTCTGGCCCC 2326
Db 481 CAGCGCGCGCGCGCGCTTCTTAACTCTCAACCTGACAGCTTTCACCTGAGTCTGGCCCC 540
QY 2327 TCCAGCGCGCGCAGGAGTGTGCGGGATGCGGATCAGCTGGGAGGAGGAGGGGTGCTT 2386
Db 541 TCCAGCGCGCAGGAGTGTGCGGGATGCGGATCAGCTGGGAGGAGGAGGGGTGCTT 600
QY 2387 CCACCCGAGGGG- AAGATGCTCTCGGACAG- TTTCCCGGGCAGCTCTCTGGCC- AGCTTC 2443
Db 601 CCACCCGAGGGGAAGATGCTCTCGGACAGATTACCCGGCAGCTCTCTGGCCGAGCTTC 660
QY 2444 CAGCGCCAGAGTCTCTCAAGTCCAGGCGACC- TTGGGCGCCAGCGCAGGAGAGATCCAGGTG 2502
Db 661 CAGCGCCAGAGTCTCTCAAGTCCAGGCGACC- TTGGGCGCCAGCGCAGGAGAGATCCAGGTG 720
QY 2503 G 2503
Db 721 G 721

RESULT 14
BI861206
LOCUS 603389971F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5399224 5',
DEFINITION mRNA sequence.
ACCESSION BI861206
VERSION BI861206.1 GI:16001953
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 727)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12017 row: m column: 17
High quality sequence stop: 708.
FEATURES
Location/Qualifiers
1..727
source

```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5399224"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_87"
/note="Organ: Breast; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN
Query Match
Best Local Similarity 18.0%; Score 664.2; DB 12; Length 727;
Matches 721; Conservative 0; Mismatches 3; Indels 5; Gaps 5;

QY 839 CGTCCAGTCCACAGAGAGGCCACAGAGGCCACGAGAGGTGCAGACCCCTGGGCCCGCAGCG 898
Db |||
QY 1 CGTCCAGTCCACAGAGAGGCCACAGAGGCCACGAGAGGTGCAGACCCCTGGGCCCGCAGCG 60
Db |||
QY 899 AGCCA-GAGACAGCCACATTGGCGCCCGGCTCTTCACAGACAGCTGTCTACTGACCCA 957
Db |||
QY 61 AGCCATGAGACAGCCACATTGGCGCCCGGCTCTTCACAGACAGCTGTCTACTGACCCA 120
QY 958 GCCCGACCCCTCTCTGGCCCGCCAGCTGGCAGCGAGAACGGCCAGAGCTGCACAGC 1017
Db |||
QY 121 GCCCGACCCCTCTCTGGCCCGCCAGCTGGCAGCGAGAACGGCCAGAGCTGCACAGC 180
QY 1018 AGCAGCACAGGCCA-GAGCCAGAGCCACGAGGGGACCCCAAGGAGGAGGAGGTGC 1076
Db |||
QY 181 AGCAGCACAGGTCATGAGCCAGAGCCACGAGGGGACCCCAAGGAGGAGGAGGTGC 240
QY 1077 TGCACCCACCATGTGGCTGGGAGCCAGAACGGCTGGCTCTATGTGCACTGGCTGTGGC 1136
Db |||
QY 241 TGCACCCACCATGTGGCTGGGAGCCAGAACGGCTGGCTCTATGTGCACTGGCTGTGGC 300
QY 1137 CAACCTGGAAGAGTGCCTGCACTCCATCAAGCTGAAGATTCTGTGCTGAGCCTGGTGCA 1196
Db |||
QY 301 CAACCTGGAAGAGTGCCTGCACTCCATCAAGCTGAAGATTCTGTGCTGAGCCTGGTGCA 360
QY 1197 TGTCAAGCGCGTGTGTGGTGTGGCTGTGGCGAGCGGACCCCTGGCCATCTTCCACCGTGG 1256
Db |||
QY 361 TGTCAAGCGCGTGTGTGGTGTGGCTGTGGCGAGCGGACCCCTGGCCATCTTCCACCGTGG 420
QY 1257 TGAAGATGGCCAGTGGGATCTGAGCACTATCACTAATGACCTGGGCCACCGCCACCA 1316
Db |||
QY 421 TGAAGATGGCCAGTGGGATCTGAGCACTATCACTAATGACCTGGGCCA-CGCGACCA 479
QY 1317 CTCATCCGCTGCATGGCTGTGTGTACGACCGGTGTGGTGTGGCTACAGAACAAAGGT 1376
Db |||
QY 480 CTCATCCGCTGCATGGCTGTGTGTACGACCGGTGTGGTGTGGCTACAGAACAAAGGT 539
QY 1377 GCAGCTATCCAGCCCAAGACCATGCGAGATAGAGAAGTCAATTTGACGCCCAAGCGCGCG 1436
Db |||
QY 540 GCAGCTATCCAGCCCAAGACCATGCGAGATAGAGAAGTCAATTTGACGCCCA-CGCGCGCG 598
QY 1437 GGAGAGCCAGTGGCGAGCTGGCTGGATGGCGATGGCGATGGGTGTTCATCCGCT 1496
Db |||
QY 599 GGAGAGCCAGTGGCGAGCTGGCTGGATGGCGATGGGTGTTCATCCGCT 658
QY 1497 GGAATCCACCTCGAGGCTCTACCATGACACAGCCACAGCATCTACAGGA-CGTGGACA 1555
Db |||
QY 659 GGAATCCACCTCGAGGCTCTACCATGACACAGCCACAGCATCTACAGGA-CGTGGACA 718
QY 1556 TTGAGCCCT 1564
Db |||
QY 719 TAGAGCCCT 727

RESULT 15
BU614626
```

LOCUS BU614626 862 bp mRNA linear EST 20-FEB-2003
DEFINITION UI-M-EVO-cbh-m-09-0-UI.r1 NIH BMAP EVO Mus musculus cDNA clone
UI-M-EVO-cbh-m-09-0-UI 5', mRNA sequence.
ACCESSION BU614626
VERSION BU614626.1 GI:23280841
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 862)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..862
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-EVO-cbh-m-09-0-UI"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP EVO"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP); 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

Query Match 17.9%; Score 660; DB 13; Length 862;
Best Local Similarity 85.5%; Pred. No. 4.5e-99;
Matches 735; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 709 GGTATCACCTGGTGGCTGTGCACCCGCTGCAACGTCGCGGAGCACTGCTCTCC 768
Db |||
QY 2 GGCATCACCTGGTGGCTGTGCACCCGCTGCAACGTCGCGGAGCACTGCTCTCA 61
Db |||
QY 769 CGAGGGAGCACCCCGTGTAGACAAGGGGCAAGGGGAGGTGGCCACCATCGCCACGGG 828
Db |||
QY 62 CGAGGAGACACCCCGTGTAGACAAGGGGCAAGGGGAGGTGGCCACCATCGG 121
Db |||
QY 829 AAGGTCAACCCGTCCTCAGTCCACAGAGAGGCCACAGAGCCACGAGGTGCAGACCT 888
Db |||
QY 122 AAGGTCAACCCGTCCTCAGTCCACAGAGAGGCCACAGAGCCACGAGGTGCAGACCT 181
Db |||
QY 889 GGGCCCGAGGAGCCAGAGAGAGCCACATTCGCGCCCGGGCTCTCAGAGCAGCTTTC 948
Db |||
QY 182 GGTCCCGAGGAGTCAGAGAGAGCCACATTCGCGCCCGGGCTCTCAGAGCAGCTTTC 241
Db |||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 00:04:58 ; Search time 13735 Seconds
(without alignments)
11657.020 Million cell updates/sec

Title: US-10-019-495-26
Perfect score: 3694
Sequence: 1 ctggcaggttcttagtgc.....gtcgtataaaaaaaaaa 3694

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb.pa.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3694	100.0	3694	6	BD269958	Human mem
2	3694	100.0	3694	6	AX041037	AX041037 Sequence
3	3484.6	94.3	4701	6	AX440473	AX440473 Sequence
4	3482	94.3	5621	9	AB028989	AB028989 Homo sapi
5	1816.2	49.2	2338	9	AB071076	AB071076 Macaca fa
6	1807	48.9	5440	10	BC068603	BC068603 Mus muscu
7	1797.4	48.7	5443	10	AF178636	AF178636 Mus muscu
8	1780.6	48.2	5411	10	AB093282	AB093282 Mus muscu
9	1770.4	47.9	5532	10	AF178637	AF178637 Mus muscu
10	1751.4	47.4	20612	9	HS371H6	AL031718 Human DNA
11	1751.4	47.4	270150	9	AE006639	AE006639 Homo sapi
12	1746.6	47.3	69437	9	AC012180	AC012180 Homo sapi
13	1542.6	41.8	4545	10	AF262046	AF262046 Mus muscu
14	1531.6	41.5	4173	10	AB005662	AB005662 Mus muscu
15	1497.6	40.5	3945	10	AB043123	AB043123 Mus muscu
16	1497.6	40.5	4011	10	AB043125	AB043125 Mus muscu
17	1497.6	40.5	4014	10	AB043124	AB043124 Mus muscu
18	520	14.1	1728	10	BC004003	BC004003 Mus muscu
19	505.2	13.7	5311	10	BC060506	BC060506 Mus muscu
20	463.2	12.5	2316	9	AF542172	AF542172 Homo sapi
21	442.2	12.0	4667	10	AF327451	AF327451 Mus muscu
22	440.6	11.9	4221	10	BC060100	BC060100 Mus muscu
23	439.4	11.9	182073	2	AC114907	AC114907 Mus muscu
24	439.4	11.9	212298	2	AC121120	AC121120 Mus muscu
25	436.2	11.8	175759	10	AF220294	AF220294 Mus muscu
26	413.4	11.2	3825	9	AB011088	AB011088 Homo sapi
27	413.4	11.2	4675	9	AF327452	AF327452 Homo sapi
28	389	10.5	199278	2	AC130925	AC130925 Rattus no
29	389	10.5	239810	2	AC111901	AC111901 Rattus no
30	370.2	10.0	3113	6	AR083288	AR083288 Sequence
31	370.2	10.0	3113	6	AX099558	AX099558 Sequence
32	370.2	10.0	3113	6	BD056704	BD056704 Secreter
33	360.6	9.8	3791	6	AX281594	AX281594 Sequence
34	352.4	9.5	43288	9	HS361A3	AL031717 Human DNA
35	294.4	8.0	4239	3	AF262045	AF262045 Drosophil
36	294.4	8.0	4455	3	BT001718	BT001718 Drosophil
37	279	7.6	1966	6	AX880664	AX880664 Sequence
38	279	7.6	1966	6	BD158516	BD158516 Primer fo
39	279	7.6	1966	9	AK023512	AK023512 Homo sapi
40	252.8	6.8	1037	3	AY060748	AY060748 Drosophil
41	244	6.6	633	6	AX869408	AX869408 Sequence
42	244	6.6	633	6	BD149470	BD149470 Primer fo
43	237	6.4	237	6	AX884437	AX884437 Sequence
44	237	6.4	237	6	BD024047	BD024047 Sequence
45	220.4	6.0	10938	2	AC014510	AC014510 Drosophil

ALIGNMENTS

RESULT 1
BD269958
LOCUS BD269958 3694 bp DNA linear PAT 17-JUL-2003
DEFINITION Human membrane-associated proteins.
ACCESSION BD269958
VERSION BD269958.1 GI:33079726
KEYWORDS JP 2002542782-A/9.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3694)
AUTHORS Hillman,J.L., Bandman,O., Tang,T.Y., Lal,P., Yue,H., Reddy,R.,
Azimzai,Y. and Baughn,M.R.
TITLE Human membrane-associated proteins

JOURNAL	Patent: JP 2002542782-A 9 17-DEC-2002; INCYTE GENOMICS INC OS Homo sapiens (human) PN JP 2002542782-A/9 PD 17-DEC-2002 PF 20-APR-2000 JP 2000614390 PR 23-APR-1999 US 60/130694, 23-JUN-1999 US 60/140580 PI JENNIFER L HILLMAN, OLGA BANDMAN, TOM Y TANG, PREETI LAL, HENRY PI YUE, PI ROOPA REDDY, VALDA AZIMZAI, MARIAH R BAUGHN PC C12N15/09, A61K31/711, A61K38/00, A61K45/00, A61K48/00, A61P1/16, PC A61P1/18, PC A61P3/10, A61P5/02, A61P5/06, A61P5/14, A61P5/26, A61P5/30, A61P7/00, PC A61P7/02, PC A61P7/06, A61P9/00, A61P9/10, A61P9/12, A61P13/00, A61P13/02, A61P13/ PC 08, PC A61P13/12, A61P15/00, A61P15/08, A61P15/12, A61P15/14, A61P17/00, PC A61P17/06, PC A61P19/00, A61P19/08, A61P19/10, A61P21/04, A61P25/08, A61P25/18, PC A61P27/00, PC A61P27/06, A61P27/12, A61P31/00, A61P35/00, A61P35/02, A61P37/02, PC A61P43/00, PC A61P43/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC 10, PC C12P21/02, C12Q1/68, C12N15/00, C12N5/00, A61K37/02 CC Incyte ID No.: 5733930CB1 FH Key Location/Qualifiers FT source 1..3694 /organism='Homo sapiens (human)'. FT Location/Qualifiers 1..3694 /organism='Homo sapiens' /mol_type='genomic DNA' /db_xref='taxon:9606'
FEATURES	source
ORIGIN	Query Match 100.0%; Score 3694; DB 6; Length 3694; Best Local Similarity 100.0%; Pred. No. 0; Matches 3694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CTGGCAGGTTCTTAGTGAGCCCTCTGGGCAGGGGTCTGGAGGTCGGAGAAGGAGGCCCG 60
DB	1 CTGGCAGGTTCTTAGTGAGCCCTCTGGGCAGGGGTCTGGAGGTCGGAGAAGGAGGCCCG 60
QY	61 CGGCTCCACGCCGCCCACTGCCTCCAGTCTTAAGTGGCTCTGCGCGCAGAACTGGCA 120
DB	61 CGGCTCCACGCCGCCCACTGCCTCCAGTCTTAAGTGGCTCTGCGCGCAGAACTGGCA 120
QY	121 GCCCTACCTGCTCAGATTTCTGGAGGGATGGGTAGGAGCCAGAGGTTTCGTCCCCACGGCG 180
DB	121 GCCCTACCTGCTCAGATTTCTGGAGGGATGGGTAGGAGCCAGAGGTTTCGTCCCCACGGCG 180
QY	181 CCTCCCTGCTCCCTGCAGCTCAGTCCCACCGGGGCCAGGAGGACA CGCGGATGAAGAAC 240
DB	181 CCTCCCTGCTCCCTGCAGCTCAGTCCCACCGGGGCCAGGAGGACA CGCGGATGAAGAAC 240
QY	241 GTGCCGGTCCGGTGTACTCGGCCCTCTGGTGAGAGGACCCACCATTGAAGCTGTGG 300
DB	241 GTGCCGGTCCGGTGTACTCGGCCCTCTGGTGAGAGGACCCACCATTGAAGCTGTGG 300
QY	301 TGTGCCGGGGCGTCAAACCTCAGCGGGGTGGAGGCCCAATGAGACGACGCTGGGAATGGA 360
DB	301 TGTGCCGGGGCGTCAAACCTCAGCGGGGTGGAGGCCCAATGAGACGACGCTGGGAATGGA 360
QY	361 GTCAAGCCAGCCAGGCGCGGATCCCTGACCTCGACCCGGGAAGGAGACGGCGAGCCC 420
DB	361 GTCAAGCCAGCCAGGCGCGGATCCCTGACCTCGACCCGGGAAGGAGACGGCGAGCCC 420
QY	421 AAGAGGCCCACACGCTCTCCCGAAGAGGAAGGCCAAGGAGCTCCCTGAAATGACGCC 480
DB	421 AAGAGGCCCACACGCTCTCCCGAAGAGGAAGGCCAAGGAGCTCCCTGAAATGACGCC 480
COMMENT	

Qy	1561	CCCTAGCTCAGCAAGATGCTAGGCATCTGGCACTGGGTTTCTCTCTCGTACGCATCACG	1620
Db	1561	CCCTAGCTCAGCAAGATGCTAGGCATCTGGCACTGGGTTTCTCTCTCGTACGCATCACG	1620
Qy	1621	GGCCTGCTTGTGCGGGGAGCGGGCTCTTGGGTGGGCACCGCAACAGGAGTGGTCACTCC	1680
Db	1621	GGCCTGCTTGTGCGGGGAGCGGGCTCTTGGGTGGGCACCGCAACAGGAGTGGTCACTCC	1680
Qy	1681	ATCCCCCTTGACAGACACTGTGTCTCTGCACCGAGCCAGCTCCTGGGGCTCCGAGCCAAAT	1740
Db	1681	ATCCCCCTTGACAGACACTGTGTCTCTGCACCGAGCCAGCTCCTGGGGCTCCGAGCCAAAT	1740
Qy	1741	AAGACATCCCCACCTCTGGGAGGGCGCCGTCCTCCGGGGGCATCATCCAGCTGTATGGC	1800
Db	1741	AAGACATCCCCACCTCTGGGAGGGCGCCGTCCTCCGGGGGCATCATCCAGCTGTATGGC	1800
Qy	1801	GATGACAGAGTGACAGGGCGGCAGCAGCTTTCATTCCTCTACTGCTCCATGGCCAGGCC	1860
Db	1801	GATGACAGAGTGACAGGGCGGCAGCAGCTTTCATTCCTCTACTGCTCCATGGCCAGGCC	1860
Qy	1861	CAGCTATCTTCCATGGGCACCGGATGTCGCTGAAGTTCCTTGTCTCGGTGCCAGGGAAC	1920
Db	1861	CAGCTATCTTCCATGGGCACCGGATGTCGCTGAAGTTCCTTGTCTCGGTGCCAGGGAAC	1920
Qy	1921	GTGCTGGCCACCCCTGAATGGCAGTGTGTGTGACAGCCAGCCGAGGGCCCTGGGCCAGCT	1980
Db	1921	GTGCTGGCCACCCCTGAATGGCAGTGTGTGTGACAGCCAGCCGAGGGCCCTGGGCCAGCT	1980
Qy	1981	GGCCCTGGCTCGGAGGTCGAGGGCCAGAAGCTGGGAAAGTGTGTGTGTGAGCGCGGG	2040
Db	1981	GGCCCTGGCTCGGAGGTCGAGGGCCAGAAGCTGGGAAAGTGTGTGTGTGAGCGCGGG	2040
Qy	2041	GAGGCTACATCGACTTCCGATTTGGAGACGGAGGACACGACGACGAGGAGGGGCGCA	2100
Db	2041	GAGGCTACATCGACTTCCGATTTGGAGACGGAGGACACGACGACGAGGAGGGGCGCA	2100
Qy	2101	GGGACATGAGCCAGGTGAAGCCCGTGTGTCTCAAGGCAGAGGGCGAGTACATCATCTGTG	2160
Db	2101	GGGACATGAGCCAGGTGAAGCCCGTGTGTCTCAAGGCAGAGGGCGAGTACATCATCTGTG	2160
Qy	2161	TGGCAGGTGTCTTACACCCCGAGTGAAGTGTGTGCTGCTGGGCCGACCTGTACATA	2220
Db	2161	TGGCAGGTGTCTTACACCCCGAGTGAAGTGTGTGCTGCTGGGCCGACCTGTACATA	2220
Qy	2221	GGACCCCGACCACTGACACCCCGCCCGGGGTAGCAGCAGCAGCGCCGCGCGG	2280
Db	2221	GGACCCCGACCACTGACACCCCGCCCGGGGTAGCAGCAGCAGCGCCGCGCGG	2280
Qy	2281	CCTCTTTAACTCTCAACTGCACTTTTCACTGAGTCTGGCCCTCCAGCGGGCAGGG	2340
Db	2281	CCTCTTTAACTCTCAACTGCACTTTTCACTGAGTCTGGCCCTCCAGCGGGCAGGG	2340
Qy	2341	AGTGGGGGTGCGGATCAGCTGGGAGGAGGGAGGGGTGCTTCCACCGAGGGGAA	2400
Db	2341	AGTGGGGGTGCGGATCAGCTGGGAGGAGGGAGGGGTGCTTCCACCGAGGGGAA	2400
Qy	2401	GATGCTCTCGGGAAGTTTCCCGGGAGCTCTTGGCCAGTTCAGAGTGGTCTTACCTAA	2460
Db	2401	GATGCTCTCGGGAAGTTTCCCGGGAGCTCTTGGCCAGTTCAGAGTGGTCTTACCTAA	2460
Qy	2461	GTCAGGGCACCTTGGGCCACGAGCGAGGAGAAATCCGAGTGGTCTTACCTGG	2520
Db	2461	GTCAGGGCACCTTGGGCCACGAGCGAGGAGAAATCCGAGTGGTCTTACCTGG	2520
Qy	2521	GCCTCTTACTCCCGAGCACCTTGGAGGAGGACAGGGGCTCCCGCCGCGAGGCTGCCTG	2580
Db	2521	GCCTCTTACTCCCGAGCACCTTGGAGGAGGACAGGGGCTCCCGCCGCGAGGCTGCCTG	2580
Qy	2581	CCCTGGGCCACCTCTGCATGCTGCTCATGGGCGACCTTGCCTCTTGGGCCCTCACTCT	2640
Db	2581	CCCTGGGCCACCTCTGCATGCTGCTCATGGGCGACCTTGCCTCTTGGGCCCTCACTCT	2640

QY	2641	GCCTAAGGGAGGTGGGCCAGAGCACTAGCCTTTGCCACGAGGAGGTGGGCCTCAGCGCTGCC	2700
DB	2641	GCTTAGGGGAGCTGGGCCAGGCACTAGCCTTTGCCACGAGGAGGTGGGCCTCAGCGCTGCC	2700
QY	2701	AGGTCCCTGCACCCCCAGCCGGCCTTCTCTGGGGCCTCCCGTGCTCAAGCCTCTATCCTG	2760
DB	2701	AGTGCCTGCACCCAGCCGGCCTTCTCTGGGGCCTCCCGTGCTCAAGCCTCTATCCTG	2760
QY	2761	TCTGTCCCACCCAGCTGTCCCCTGCCAGGAGCTGGCATAAAGCAGAGGCCCGGC	2820
DB	2761	TCTGTCCCACCCAGCTGTCCCCTGCCAGGAGCTGGCATAAAGCAGAGGCCCGGC	2820
QY	2821	TCCCTGGGGAGCTGCTTGAGAAACAGAGACTGCTTACCCCATCTGCCCCATGACAGGAGGC	2880
DB	2821	TCCCTGGGGAGCTGCTTGAGAAACAGAGACTGCTTACCCCATCTGCCCCATGACAGGAGGC	2880
QY	2881	TCTTGCACCCCGCTTCTGACCCGTGCCCCCAGAGCTCTGCTTGGGCAGAGACTCAC	2940
DB	2881	TCTTGCACCCCGCTTCTGACCCGTGCCCCCAGAGCTCTGCTTGGGCAGAGACTCAC	2940
QY	2941	TTGAGGAGTGGGCCCTTGGAGTCTGTCTCCCTCCAGAAAGCCACAGGCTGGGATTTCTCA	3000
DB	2941	TTGAGGAGTGGGCCCTTGGAGTCTGTCTCCCTCCAGAAAGCCACAGGCTGGGATTTCTCA	3000
QY	3001	GGCTGCACGGCAGGCCCTCAGGCTCAGGAAGAGGGGAGGCCCTTGGCTCTCCGGGATC	3060
DB	3001	GGCTGCACGGCAGGCCCTCAGGAAGAGGGGAGGCCCTTGGCTCTCCGGGATC	3060
QY	3061	AGTCTTAGGACACAGGCTCAGCCTCAGTTGATGGGGGATGATGTCTCCCGGGCCTGC	3120
DB	3061	AGTCTTAGGACACAGGCTCAGCCTCAGTTGATGGGGGATGATGTCTCCCGGGCCTGC	3120
QY	3121	CTCTTGACCGGGCTCCACGGAGCCAGCTTCCAGACACGCTACTAAAGTGCCTAGGGTTG	3180
DB	3121	CTCTTGACCGGGCTCCACGGAGCCAGCTTCCAGACACGCTACTAAAGTGCCTAGGGTTG	3180
QY	3181	CCCCTGTGGCTGTCTCCACGGAGCAACAGAGAGCCACCAAGCAGAGGCCCTTGGGCG	3240
DB	3181	CCCCTGTGGCTGTCTCCACGGAGCAACAGAGAGCCACCAAGCAGAGGCCCTTGGGCG	3240
QY	3241	TGAGGATGGAGCCGCCCCAGCCGACTCCAGCCCGCAGAGGCAGAGCCACCTTGGAC	3300
DB	3241	TGAGGATGGAGCCGCCCCAGCCGACTCCAGCCCGCAGAGGCAGAGCCACCTTGGAC	3300
QY	3301	TGCTCTCCCTGCCAGCTGGGCTCTCTGGCCTATTCTTACCTTCCAGGCCACTGCACT	3360
DB	3301	TGCTCTCCCTGCCAGCTGGGCTCTCTGGCCTATTCTTACCTTCCAGGCCACTGCACT	3360
QY	3361	CCTGTCTGGAGCCCTTATGAGGGCAGCCACGCCCCCGCACCCACCCCAACACAGAA	3420
DB	3361	CCTGTCTGGAGCCCTTATGAGGGCAGCCACGCCCCCGCACCCACCCCAACACAGAA	3420
QY	3421	GCCAGATCTTGGGAGCTGCCGCCAAGCCCGCTGGCCACCGAGGCTTGACGCGCTG	3480
DB	3421	GCCAGATCTTGGGAGCTGCCGCCAAGCCCGCTGGCCACCGAGGCTTGACGCGCTG	3480
QY	3481	CGCTGCGGCTTCTCCCAACCCCTGCCACTTCCACTTGATGTATGTCGCTCCCTCG	3540
DB	3481	CGCTGCGGCTTCTCCCAACCCCTGCCACTTCCACTTGATGTATGTCGCTCCCTCG	3540
QY	3541	TCTGTTCCCCAGGATCTCGAAGTGACTCCGGGCTGAGCAGTGGGGCGGCTGGGGGAGGG	3600
DB	3541	TCTGTTCCCCAGGATCTCGAAGTGACTCCGGGCTGAGCAGTGGGGCGGCTGGGGGAGGG	3600
QY	3601	GTGACGATCTCTCAGGCTTTGGCCCTGCAAGCAAAACCCATATCTGCTCTGTATGTA	3660
DB	3601	GTGACGATCTCTCAGGCTTTGGCCCTGCAAGCAAAACCCATATCTGCTCTGTATGTA	3660
QY	3661	ATAAATGCTTAAAGCTGTAAAAAATAAAAAA	3694
DB	3661	ATAAATGCTTAAAGCTGTAAAAAATAAAAAA	3694

RESULT 2
AX041037
LOCUS
DEFINITION
Sequence 26 from Patent WO0065054.
AX041037
ACCESSION
AX041037.1 GI:11340615
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Hillman, J.L., Bandman, O., Tang, Y.T., Lal, P., Yue, H., Redd, R.,
Azimzai, Y. and Baughn, M.R.
Human membrane-associated proteins
Patent: WO 0065054-A 26 02-NOV-2000;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..3694
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No.: 5733930CB1"

ORIGIN
Query Match 100.0%; Score 3694; DB 6; Length 3694;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGGCAGGTTCTAGTAGACCTCTCGGCGAGGGGTCTGGAGGTCGGAGAGGAGGCCCG 60
DB 1 CTGGCAGGTTCTAGTAGACCTCTCTGGCAGGGGTCTGGAGGTCGGAGGAGGCCCG 60
QY 61 CGGCTCCACAGCCCCACCTGCTCCAGTCTTGAAGTGGTCTGCGCCAGAAAGTGGCA 120
DB 61 CGGCTCCACAGCCCCACCTGCTCCAGTCTTGAAGTGGTCTGCGCCAGAAAGTGGCA 120
QY 121 GCCCTACCTGCTCAGATTCTCGAGGATGGGTAGGACCCAGGTTCTGTCGCCACGGG 180
DB 121 GCCCTACCTGCTCAGATTCTCGAGGATGGGTAGGACCCAGGTTCTGTCGCCACGGG 180
QY 181 CCTCCCTGCTCCCTCAGCTGAGTCCCAACGGGGCCAGGAGCACGGGATGAAGAAC 240
DB 181 CCTCCCTGCTCCCTCAGCTGAGTCCCAACGGGGCCAGGAGCACGGGATGAAGAAC 240
QY 241 GTGCGGTGTCGGGTGTAAGTGGTCTGCTGGTGGAGAGGACCCCAACATGAAGTGTGG 300
DB 241 GTGCGGTGTCGGGTGTAAGTGGTCTGCTGGTGGAGAGGACCCCAACATGAAGTGTGG 300
QY 301 TGTGCGGGGGCTCAACTGAGCGGTGGAGCCCAATGAGACGACGCTGGGATGGA 360
DB 301 TGTGCGGGGGCTCAACTGAGCGGTGGAGCCCAATGAGACGACGCTGGGATGGA 360
QY 361 GTCAAGCCACGCGCGAGCCGCGCATCCCTGAGCTGCGACCGGAAAGAGACGCGAGCCC 420
DB 361 GTCAAGCCACGCGCGAGCCGCGCATCCCTGAGCTGCGACCGGAAAGAGACGCGAGCCC 420
QY 421 AAGAGCGCCCAACAGTCTCCGAGAGAAAGAGGCGCAAGAGGCTCCCTGAATGAGACGCC 480
DB 421 AAGAGCGCCCAACAGTCTCCGAGAGAAAGAGGCGCAAGAGGCTCCCTGAATGAGACGCC 480
QY 481 ACTCTCAGCGGGTGTGGATCTGACGACGCTGACCAACGACGAGGTGGTATCATC 540
DB 481 ACTCTCAGCGGGTGTGGATCTGACGACGCTGACCAACGACGAGGTGGTATCATC 540
QY 541 GACGCCAACACGCGCGGACGCTGGTGGACCAAGTTACCGCTCTGCAACGCGACGCTGCTG 600
DB 541 GACGCCAACACGCGCGGACGCTGGTGGACCAAGTTACCGCTCTGCAACGCGACGCTGCTG 600
QY 601 TGCATCTCAGCATCTCCCGGGGCGAGGACGAGCACTACCTCCCGGGGAGATGTTCTTG 660
DB 601 TGCATCTCAGCATCTCCCGGGGCGAGGACGAGCACTACCTCCCGGGGAGATGTTCTTG 660

QY 661 GACAGCAGCTGAACCCAGAGGACCCGCGCGCAGATGGCGTGTGGCGGTATCACCGTG 720
DB 661 GACAGCAGCTGAACCCAGAGGACCCGCGCGCAGATGGCGTGTGGCGGTATCACCGTG 720
QY 721 GTGGGCTGTGCCACCCCGCTGCAACGTCGCGGAGCAACTGCTCTCTCCGAGGGGACACC 780
DB 721 GTGGGCTGTGCCACCCCGCTGCAACGTCGCGGAGCAACTGCTCTCTCCGAGGGGACACC 780
QY 781 CCAGTGTAGACAGGGGCGAGGGAGGTGGCCACCATCGCCACGAGGAGGTCAACCCG 840
DB 781 CCAGTGTAGACAGGGGCGAGGGAGGTGGCCACCATCGCCACGAGGAGGTCAACCCG 840
QY 841 TCCAGTGTAGACAGGGGCGAGGGAGGTGGCCACCATCGCCACGAGGAGGTCAACCCG 900
DB 841 TCCAGTGTAGACAGGGGCGAGGGAGGTGGCCACCATCGCCACGAGGAGGTCAACCCG 900
QY 901 CCAGAGACAGCCACATTTGGCGCCCGGGCTCTCAGAGACAGCTTCTCAGTGAACCCAGCC 960
DB 901 CCAGAGACAGCCACATTTGGCGCCCGGGCTCTCAGAGACAGCTTCTCAGTGAACCCAGCC 960
QY 961 CCGACCCCGTCTCTGGCCCCCAGCCTGGCAGCGAGAACGGGCCAGAGCCTGACAGCAGC 1020
DB 961 CCGACCCCGTCTCTGGCCCCCAGCCTGGCAGCGAGAACGGGCCAGAGCCTGACAGCAGC 1020
QY 1021 AGCACAGCGCCAGAGCCAGAGCCCGGGGACCCACAGGAGAGGAGGAGTGTCTGCA 1080
DB 1021 AGCACAGCGCCAGAGCCAGAGCCCGGGGACCCACAGGAGAGGAGGAGTGTCTGCA 1080
QY 1081 CCCACCATGTGGCTGGGAGCCAGAACCGCTGGCTCTAATGTGCACTCGGCTGTGGCCAAC 1140
DB 1081 CCCACCATGTGGCTGGGAGCCAGAACCGCTGGCTCTAATGTGCACTCGGCTGTGGCCAAC 1140
QY 1141 TGGAGAGTGCCTGCATCCATCAAGCTGAAGGATTTCTGTGCTGAGCCTGTGTGATGTC 1200
DB 1141 TGGAGAGTGCCTGCATCCATCAAGCTGAAGGATTTCTGTGCTGAGCCTGTGTGATGTC 1200
QY 1201 AAAGCCCGTGTGTGGTGGCTCTGGCGGACCGGACCCCTGGCCATTTCCACCGTGGTGA 1260
DB 1201 AAAGCCCGTGTGTGGTGGCTCTGGCGGACCGGACCCCTGGCCATTTCCACCGTGGTGA 1260
QY 1261 GATGCCAGTGGGATCTGAGCAATATCACCTAATGAGACCTGGGCGACCCGACCACTCC 1320
DB 1261 GATGCCAGTGGGATCTGAGCAATATCACCTAATGAGACCTGGGCGACCCGACCACTCC 1320
QY 1321 ATCCGCTGCATGGCTGTGTGACACCGCTGTGCTGAGTGTGCTCAAGAACAAAGTGTGAC 1380
DB 1321 ATCCGCTGCATGGCTGTGTGACACCGCTGTGCTGAGTGTGCTCAAGAACAAAGTGTGAC 1380
QY 1381 GTCATCCAGCCCAAGACCATGCAATAGAGAGTCAATTTGACGCCCAACCCGCGGGGAG 1440
DB 1381 GTCATCCAGCCCAAGACCATGCAATAGAGAGTCAATTTGACGCCCAACCCGCGGGGAG 1440
QY 1441 AGCCAGGTGCGGACGCTGGTGGATCGGCGATGGCGTATGGGTTCATCCGCTGGAC 1500
DB 1441 AGCCAGGTGCGGACGCTGGTGGATCGGCGATGGCGTATGGGTTCATCCGCTGGAC 1500
QY 1501 TCCACCTGAGGCTCTACCATGCACACGACGACGACGACGACGACGACGACGACGACGAC 1560
DB 1501 TCCACCTGAGGCTCTACCATGCACACGACGACGACGACGACGACGACGACGACGACGAC 1560
QY 1561 CCCTACGTCAGCAAGATGCTAGGCACTGCAAGCTGGGTTCTCTTGTGTAGCATCAG 1620
DB 1561 CCCTACGTCAGCAAGATGCTAGGCACTGCAAGCTGGGTTCTCTTGTGTAGCATCAG 1620
QY 1621 GCCCTGCTGTGCGGGGACGCGGCTCTGGGTGGGACCGGCAACGGAGTGTCTCTCC 1680
DB 1621 GCCCTGCTGTGCGGGGACGCGGCTCTGGGTGGGACCGGCAACGGAGTGTCTCTCC 1680
QY 1681 ATCCCCCTGACAGAGACTGTGTCTGTGACCGAGGCGAGCTCTCTGGGGTCCGAGCAAT 1740
DB 1681 ATCCCCCTGACAGAGACTGTGTCTGTGACCGAGGCGAGCTCTCTGGGGTCCGAGCAAT 1740
QY 1741 AAGACATCCCCCAACCTCTGGGGAGGGCGCCCTCCCGGGGGGCAATCATCCAGTGTATGC 1800

1741 AAGACATCCCACTCTCGGGAGGCGCCGTCCTCCGGGGCATCATCACTGATATGGC 1800
1801 GATGACAGAGTACAGAGGGGGCCAGCAGTTTCATCCCTACTGTCTCCATGAGCCAGGCG 1860
1801 GATGACAGAGTACAGAGGGGGCCAGCAGTTTCATCCCTACTGTCTCCATGAGCCAGGCG 1860
1861 CAGCTATGCTTCCATGGGACACCGCATGCCGTGGAAGTTCTTTGTCTCGGTGCCAGGGAAC 1920
1861 CAGCTATGCTTCCATGGGACACCGCATGCCGTGGAAGTTCTTTGTCTCGGTGCCAGGGAAC 1920
1921 GTGCTGGCCACCTGAAATGGAGTGTGTGGACAGCCACGCCAGAGGGCCCTGGGCCAGCT 1980
1921 GTGCTGGCCACCTGAAATGGAGTGTGTGGACAGCCACGCCAGAGGGCCCTGGGCCAGCT 1980
1981 GCCCTGCTCGAGAGTCCAGAGCCAGAGCTGCCGAACTGTCGAGAGGGCCCTGGGCCAGCT 2040
1981 GCCCTGCTCGAGAGTCCAGAGCCAGAGCTGCCGAACTGTCGAGAGGGCCCTGGGCCAGCT 2040
2041 GAGGGCTACATCGACTTCCGCAATTGGAGACGAGAGAGCAGCAGAGAGGAGGAGGCGCA 2100
2041 GAGGGCTACATCGACTTCCGCAATTGGAGACGAGAGAGCAGCAGAGAGGAGGAGGCGCA 2100
2101 GGGGACATGAGCCAGGTGAAGCCCGTGTGTCCAAAGGACAGCGCAGTCACATCATCTGTG 2160
2101 GGGGACATGAGCCAGGTGAAGCCCGTGTGTCCAAAGGACAGCGCAGTCACATCATCTGTG 2160
2161 TGGCAGGTGTCTACACCCCGAGTGAGCTGCTGCCCTGCTGAGAGGGCCCTGGGCCAGCT 2220
2161 TGGCAGGTGTCTACACCCCGAGTGAGCTGCTGCCCTGCTGAGAGGGCCCTGGGCCAGCT 2220
2221 GGAACCCCGACCACTGACCCCGCCGCGCGGGGTAGCAGCAGAGGCGCCCGCCG 2280
2221 GGAACCCCGACCACTGACCCCGCCGCGCGGGGTAGCAGCAGAGGCGCCCGCCG 2280
2281 CCTCTTAACTCTCAACTGTCAGCTTTCACTGAGTGTGCCCCCTCCAGCGGCGAGG 2340
2281 CCTCTTAACTCTCAACTGTCAGCTTTCACTGAGTGTGCCCCCTCCAGCGGCGAGG 2340
2341 AGTGGGGGATGCGATCAGCTGGAGAGGAGGAGGGGGTGTTCACCCGAGGGGAA 2400
2341 AGTGGGGGATGCGATCAGCTGGAGAGGAGGAGGGGGTGTTCACCCGAGGGGAA 2400
2401 GATGCTCTCGGACAGTTTCCCGGACAGCTCTCGCCAGCTTCCAGCCAGAGTCTCA 2460
2401 GATGCTCTCGGACAGTTTCCCGGACAGCTCTCGCCAGCTTCCAGCCAGAGTCTCA 2460
2461 GTCCAGGGACCTTGGGCGCCAGCGCAGGAGAAATCCGAGGTGGTCTTGCTCTACCTTG 2520
2461 GTCCAGGGACCTTGGGCGCCAGCGCAGGAGAAATCCGAGGTGGTCTTGCTCTACCTTG 2520
2521 GCCTCTACTTCCAGCACCTTGAGAGGAGGAGGGGCTCCCGCGCGAGGCTGCTG 2580
2521 GCCTCTACTTCCAGCACCTTGAGAGGAGGAGGGGCTCCCGCGCGAGGCTGCTG 2580
2581 CCCTGGGCCCCACCTTGATGCTGCTCATGAGGGCCACCTGCTCTGCGCCCTCACTCT 2640
2581 CCCTGGGCCCCACCTTGATGCTGCTCATGAGGGCCACCTGCTCTGCGCCCTCACTCT 2640
2641 GCCTAGGGAGCTGGGCGCAGGCACTAGCCTTTGGCCAGGAGGTGGGCTCAGGCTGCC 2700
2641 GCCTAGGGAGCTGGGCGCAGGCACTAGCCTTTGGCCAGGAGGTGGGCTCAGGCTGCC 2700
2701 AGGTGCTTGCACCCAGCGGCTTCTCTGGGGCTCCCTGCTGCTCAAGCTCTATCTG 2760
2701 AGGTGCTTGCACCCAGCGGCTTCTCTGGGGCTCCCTGCTGCTCAAGCTCTATCTG 2760
2761 TCTGTCCCCACCCAGCTGTCCCTGCGCCAGGAGCTGGCATAAAGCAGGAGCCCGCG 2820
2761 TCTGTCCCCACCCAGCTGTCCCTGCGCCAGGAGCTGGCATAAAGCAGGAGCCCGCG 2820
2821 TCCCTGGGAGCTGCTTGAGAACAGAGACTGTACCCCATCTGCCCATCAGGAGGCG 2880

2821 TCCCTGGGAGCTGCTTGAGAACAGAGACTGCTACCCCATCTGCCCATGACGAGGCGC 2880
2881 TTTTGCCAGCCCCGTTCTGACCCGTTGCCCGCAGGCTTGTGCTGGGCGAGAACTCACC 2940
2881 TTTTGCCAGCCCCGTTCTGACCCGTTGCCCGCAGGCTTGTGCTGGGCGAGAACTCACC 2940
2941 TTGGAGGAGTGGGCTTGGAGTCTGTCTCCAGAGAGCCGCCAGGGTGGGATTTCTCA 3000
2941 TTGGAGGAGTGGGCTTGGAGTCTGTCTCCAGAGAGCCGCCAGGGTGGGATTTCTCA 3000
3001 GGCTGCCAGGGCAGGCCACAGGCTCAGGAAGAGGGAGGCCCTGGGCTTCCGGGATC 3060
3001 GGCTGCCAGGGCAGGCCACAGGCTCAGGAAGAGGGAGGCCCTGGGCTTCCGGGATC 3060
3061 AGTCTTAGGACACAGGCTCAGCTCAGGTTGATGGGGATGATGTGCTCCCGGGCTG 3120
3061 AGTCTTAGGACACAGGCTCAGCTCAGGTTGATGGGGATGATGTGCTCCCGGGCTG 3120
3121 CTCCTGACGGGGCTCCACGGAGCCAGCTCCAGACAGCTACTAAGTGCCTAGGGTTG 3180
3121 CTCCTGACGGGGCTCCACGGAGCCAGCTCCAGACAGCTACTAAGTGCCTAGGGTTG 3180
3181 CCGCTGTGGGCTGCTCCAGGGAGCAACAGAGAGGCCAACAGAGGAGGCCGCTGGGGC 3240
3181 CCGCTGTGGGCTGCTCCAGGGAGCAACAGAGAGGCCAACAGAGGAGGCCGCTGGGGC 3240
3241 TGAGGATGGAGCGGCCCGCCAGCGACTCCAAAGCCGCGAGAGGCGAGACGCCACCTGGAC 3300
3241 TGAGGATGGAGCGGCCCGCCAGCGACTCCAAAGCCGCGAGAGGCGAGACGCCACCTGGAC 3300
3301 TGCTCTCCCTGCGGAGCTGGGCTCTCTGGGCTATTTCCTACCTTCAGGCGCTGCACT 3360
3301 TGCTCTCCCTGCGGAGCTGGGCTCTCTGGGCTATTTCCTACCTTCAGGCGCTGCACT 3360
3361 CCTGTCTGGGAGGCGCTTATGAGGGAGCCAGCCAGCCCGCGACCCACCCACACAGAGAA 3420
3361 CCTGTCTGGGAGGCGCTTATGAGGGAGCCAGCCAGCCCGCGACCCACCCACACAGAGAA 3420
3421 GCACAGATCTTTGGGAGCTGCCCAAGCCCGCTGGCCACCCAGGGGCTGCAGCGCTG 3480
3421 GCACAGATCTTTGGGAGCTGCCCAAGCCCGCTGGCCACCCAGGGGCTGCAGCGCTG 3480
3481 CGCTGCGGCTTCTCCCAACACCTGCACTGCACTGATGATGATGCTGCTGCTGCTGCTG 3540
3481 CGCTGCGGCTTCTCCCAACACCTGCACTGCACTGATGATGATGCTGCTGCTGCTGCTG 3540
3541 TCTGTTCGCCAGGATCTCGAAGTGACTCCGGGCTGAGCAGTGGGGGCTGGGGGAGG 3600
3541 TCTGTTCGCCAGGATCTCGAAGTGACTCCGGGCTGAGCAGTGGGGGCTGGGGGAGG 3600
3601 GTGAGATTTCTCCTCAGGCTTTGGCCCTGCAAGCAAAACCCACATATCTGCTGATGTA 3660
3601 GTGAGATTTCTCCTCAGGCTTTGGCCCTGCAAGCAAAACCCACATATCTGCTGATGTA 3660
3661 ATAAATGCTTAAAGTGTATAAAAAA 3694
3661 ATAAATGCTTAAAGTGTATAAAAAA 3694

RESULT 3

AX440473
LOCUS AX440473 4701 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 326 from Patent WO0190154.
ACCESSION AX440473
VERSION AX440473.1 GI:21665283
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Xu, J., Mitcham, J.L., Harlocker, S.L., Dillon, D.C., Secrist, H.,
Lodes, M.J., Algate, P.A., Fling, S.P., Mannion, J., Benson, D.R. and

Carter, D. Compositions and methods for the therapy and diagnosis of ovarian cancer JOURNAL Patent: WO 0190154-A 326 29-NOV-2001; CORIXA CORPORATION (US) FEATURES Location/Qualifiers source 1..4701 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"									
ORIGIN									
Query Match 94.3%; Score 3484.6; DB 6; Length 4701; Best Local Similarity 99.7%; Pred. No. 0; Matches 3490; Conservative 0; Mismatches 9; Indels 0; Gaps 0;									
QY	180	GCCTCCCTGCTCCCTGCACTGAGTCCCAACGGGGGCCAGGAGGACACGCGATGAAGAA	239						
DB	1203	GCCGCCAAGTACAAAGCAGCTGAGTCCCAACGGGGGCCAGGAGGACACGCGATGAAGAA	1262						
QY	240	CGTGCGGTGCGGTGTAAGTCCGCGCTCTGGTGGAGAGACCCACCATGAGCTGTG	299						
DB	1263	CGTGCGGTGCGGTGTAAGTCCGCGCTCTGGTGGAGAGACCCACCATGAGCTGTG	1322						
QY	300	GTGTGCGCGGGCGTCAACCTGAGCGGTGGAGGCCCAATGAGGACGACGCTGGGAATGG	359						
DB	1323	GTGTGCGCGGGCGTCAACCTGAGCGGTGGAGGCCCAATGAGGACGACGCTGGGAATGG	1382						
QY	360	AGTCAAGCCAGCGCCAGCGCGATCCCTGACCTGCGACCGCGAAGGAGCGCGAGCC	419						
DB	1383	AGTCAAGCCAGCGCCAGCGCGATCCCTGACCTGCGACCGCGAAGGAGCGCGAGCC	1442						
QY	420	CAAGAGCGCCACAGCTCTCCGAGAAAGAGGCGAAGGAGCTCCCTGAATGAGCGC	479						
DB	1443	CAAGAGCGCCACAGCTCTCCGAGAAAGGAGGAGGAGCTCCCTGAATGAGCGC	1502						
QY	480	CACCTCCAGCGGGGTGATCTCACCAGACACCTGACACAGCAGGAGGTGGTGCATCAT	539						
DB	1503	CACCTCCAGCGGGGTGATCTCACCAGACACCTGACACAGCAGGAGGTGGTGCATCAT	1562						
QY	540	CGACGCCAACACCGCGGCAAGTGGTGGACCAAGTTCACCGTCTGCAACGCGCACGTGCT	599						
DB	1563	CGACGCCAACACCGCGGCAAGTGGTGGACCAAGTTCACCGTCTGCAACGCGCACGTGCT	1622						
QY	600	GTGCATCTCCAGCATCCCGCGGCGAGGACGACGACGACTACCTCCCGGGGAGATGTTCT	659						
DB	1623	GTGCATCTCCAGCATCCCGCGGCGAGGACGACGACGACTACCTCCCGGGGAGATGTTCT	1682						
QY	660	GGACAGCGACGTAACCCAGAGGACCCGGGCGCAGATGCGTGTGCGCGGTATCACCTCT	719						
DB	1683	GGACAGCGACGTAACCCAGAGGACCCGGGCGCAGATGCGTGTGCGCGGTATCACCTCT	1742						
QY	720	GGTGGGCTGTGCCACCCGCTGCAACGTGCGCGGAGCAACTGCTCTCCCGAGGGGACAC	779						
DB	1743	GGTGGGCTGTGCCACCCGCTGCAACGTGCGCGGAGCAACTGCTCTCCCGAGGGGACAC	1802						
QY	780	CCAGTGTGTAGACAGGGGCGAGGGAGGTGGCCACCATCGCCCAAGCGGAGGTCAACCC	839						
DB	1803	CCAGTGTGTAGACAGGGGCGAGGGAGGTGGCCACCATCGCCCAAGCGGAGGTCAACCC	1862						
QY	840	GTCCAGTCCACAGAGGAGGCGCACAGAGGCCACGAGGTGCGCAGACCTTGGGCGCCAGCGA	899						
DB	1863	GTCCAGTCCACAGAGGAGGCGCACAGAGGCCACGAGGTGCGCAGACCTTGGGCGCCAGCGA	1922						
QY	900	GCCAGAGACGCCACATTGGGCGCGGGCTCTTCAACAGAGCAGTCTTCACTGACCCAGC	959						
DB	1923	GCCAGAGACGCCACATTGGGCGCGGGCTCTTCAACAGAGCAGTCTTCACTGACCCAGC	1982						
QY	960	CCCGACCCCGTCTCTTGGCCCCCAGCTCGCAGCGAGAACGGGCCAGAGCTTACAGCAG	1019						
DB	1983	CCCGACCCCGTCTCTTGGCCCCCAGCTCGCAGCGAGAACGGGCCAGAGCTTACAGCAG	2042						
QY	1020	CAGCACACGCGCCAGAGCGCCAGCGGGGACCCCAACGCGGAGCAGGACGAGTGTCTGC	1079						

DB	2043	CAGCACACGCGCCAGAGCCAGAGCCAGCGGGACCCACAGCGGAGCAGGACAGTGTGC	2102						
QY	1080	ACCCACCATGTGGCTGGAGCCAGAACGGTGTGCTCTATGTGCACTGGCTGTGGCCAA	1139						
DB	2103	ACCCACCATGTGGCTGGAGCCAGAACGGTGTGCTCTATGTGCACTGGCTGTGGCCAA	2162						
QY	1140	CTGGAAGAAGTGCCTGCACTCCATCAAGCTGAAGATTTCTGTGTGAGCCCTGGTGCATGT	1199						
DB	2163	CTGGAAGAAGTGCCTGCACTCCATCAAGCTGAAGATTTCTGTGTGAGCCCTGGTGCATGT	2222						
QY	1200	CAAAAGCCGTGTGCTGGTGGCTCTGGGCGACGGGACCCCTGGCCCATTTTCCACCGTGTGA	1259						
DB	2223	CAAAAGCCGTGTGCTGGTGGCTCTGGGCGACGGGACCCCTGGCCCATTTTCCACCGTGTGA	2282						
QY	1260	AGATGGCCAGTGGGATCTGAGCAACTATCATCTATGAACTGGGCGCACCCGACCACTC	1319						
DB	2283	AGATGGCCAGTGGGATCTGAGCAACTATCATCTATGAACTGGGCGCACCCGACCACTC	2342						
QY	1320	CATCCGCTGCATGGCTGTGTGTACGACCGGTGTGGTGTGGCTACAAAGAACAGGTGCA	1379						
DB	2343	CATCCGCTGCATGGCTGTGTGTACGACCGGTGTGGTGTGGCTACAAAGAACAGGTGCA	2402						
QY	1380	CGTCATCCAGCCCAAGACCATGCGATGAGAGATTTTGAAGCCCAACCCGCGCGGGA	1439						
DB	2403	CGTCATCCAGCCCAAGACCATGCGATGAGAGATTTTGAAGCCCAACCCGCGCGGGA	2462						
QY	1440	GAGCCAGTGGCGGAGCTGGCTGGATCGGCGATGGCGTATGGGTGTCCATCCGCTGGA	1499						
DB	2463	GAGCCAGTGGCGGAGCTGGCTGGATCGGCGATGGCGTATGGGTGTCCATCCGCTGGA	2522						
QY	1500	CTCCACCTGAGGCTCTACCATGCACACACGACCCAGCATCTTACAGGACGTGGACATTGA	1559						
DB	2523	CTCCACCTGAGGCTCTACCATGCACACACGACCCAGCATCTTACAGGACGTGGACATTGA	2582						
QY	1560	GCCTTACGTGCAAGATGCTAGGACCTGGCAAGCTGGGTTTCTCTTGTGATGATCATC	1619						
DB	2583	GCCTTACGTGCAAGATGCTAGGACCTGGCAAGCTGGGTTTCTCTTGTGATGATCATC	2642						
QY	1620	GGCCCTGTGCTGGCGGCGACCGGCTCTGGTGGGCGACCGGCAACGAGAGTGTCTCTC	1679						
DB	2643	GGCCCTGTGCTGGCGGCGACCGGCTCTGGTGGGCGACCGGCAACGAGAGTGTCTCTC	2702						
QY	1680	CATCCCCCTGCACAGAGACTGTGGTCTCTGCACCGAGGCCAGCTCTCTGGGCTCCGAGCCAA	1739						
DB	2703	CATCCCCCTGCACAGAGACTGTGGTCTCTGCACCGAGGCCAGCTCTCTGGGCTCCGAGCCAA	2762						
QY	1740	TAAGACATCCCCCACTCTGGGGAGGGCGCCCGTCCCGGGGCGATCATCCAGTGTATGG	1799						
DB	2763	TAAGACATCCCCCACTCTGGGGAGGGCGCCCGTCCCGGGGCGATCATCCAGTGTATGG	2822						
QY	1800	CGATGACAGCAGTGCAGGGCGCCAGCAGCTTCATCCCTACTGCTCCATGGCCCGAGGC	1859						
DB	2823	CGATGACAGCAGTGCAGGGCGCCAGCAGCTTCATCCCTACTGCTCCATGGCCCGAGGC	2882						
QY	1860	CCAGCTATGCTTCCATGGGCAACCGGATGCGCTGAAGTTCTTTGTCTCGGTGCCAGGAA	1919						
DB	2883	CCAGCTATGCTTCCATGGGCAACCGGATGCGCTGAAGTTCTTTGTCTCGGTGCCAGGAA	2942						
QY	1920	CGTGTGGGCAACCTTGAATGGCAGTGTGTGACAGCCCGAGCCGAGGGCCCTGGGCGAG	1979						
DB	2943	CGTGTGGGCAACCTTGAATGGCAGTGTGTGACAGCCCGAGCCGAGGGCCCTGGGCGAG	3002						
QY	1980	TGCCCCCTGCTCGAGGTCGAGGGCCAGAAAGCTGGGGAACGTGCTGGTGTGAGCGGCGG	2039						
DB	3003	TGCCCCCTGCTCGAGGTCGAGGGCCAGAAAGCTGGGGAACGTGCTGGTGTGAGCGGCGG	3062						
QY	2040	GGAGGGCTACATCGATTCGCTTCCGATTTGGAGACGGAGGAGGACGACGAGAGGAGGCGC	2099						
DB	3063	GGAGGGCTACATCGATTCGCTTCCGATTTGGAGACGGAGGAGGACGACGAGAGGAGGCGC	3122						
QY	2100	AGGGGACATGAGCCAGGTGAAGCCCGTGTGTCTCAAGGCGAGAGCGCAGTCAATCATCGT	2159						

CDS

<91...4131

/gene="KIAA1066"

/notes="This sequence was replaced that of hJ05363b cDNA as a representative cDNA sequence for KIAA1066."

/codon_start=1

/product="KIAA1066 protein"

/protein_id="BAA83018.2"

/db_xref="GI:27529814"

/translation="RAALAAVAANMTEIMDEGGGVVVYDDYCSGVMGERSVLGAG
SYREPERLLHCYDEVEEMPLVNVVLENLSDVLSNQHSEHVELELLEDNEQLLT
YERERAKQAQAEKIFEDBALDQEKLEQLQVHEFYFQTRLELAKKAYQASIGRL
EERSENKKYNAALHQHMETQTQVYHEIRSKMQVGGNSQTESSLPKARCKRPTSL
NVFLDPLQTVRAIGTQVAPAGDWHMLSDLGLOQSSSYQCPQDEMSQSGAAATP
STTCTKNTGTQSGVAAVTPNLSQPLQDYGVGSKKAKREKRSRNMVEVYQTE
MRNVYSICMGSSDSDSDVQDIIIDSTPELMCPTRIDRTGSGSTQGIYNKAFGINTDLS
YHELSTAGSVIGDVEGADLIGFESVRDDPFGMCKEYGNLLNLSNLTETKANALNV
KNDLIAKQVLSQGEVLVRLGELEAAQAKVLENIILKELEELKRVKSEALIAAREPK
EASDESVSYLTEDSKIPWAKRQRTVRVEMAKVPSYVNIEMRQYKERMELQEAQVTRTEMLR
ASREHSPVQBEKSTKIWFPSRLFSSSSPPAKPKYPSYVNIEMRQYKERMELQEAQVTRTEMLR
KMLCISAGSRPLGFPDDQTSARERKQEGYQVRHVRNDRDGLQACWSPGLAQ
AQPSGQSGQDTRMKNVVPVYCRPLVEKQDPTMKLCAAGVNLISGWRPNEDDAGVKY
PAPGRDPLTCDRREGCDPKVASTSPPEKKAKELPEMDATSGSRVILSTLTSTKVVILI
DANQPGITVDQGEVCAHVLKSTISIPASDSDYPGCEMFLSDVNVEDPGADGVLAGI
TLVGACATRVNPNRSCSRGDTPLVDKQGEVATITANGKVPNSQTEATEATEVPDPTG
AGSSAAPTMLMGAGMLYVHSGVAANAKKLSHSIKLKDSVLSLVHVKRVLVALADGT
LAIPTHRGEGDSQVQLNLMIDGPHHSIRCMNAVYIRLTHVIAHHTHQDVIDEYPSKMLQLE
KSFDAFHREGDSQVQLNLMIDGPHHSIRCMNAVYIRLTHVIAHHTHQDVIDEYPSKMLQLE
TKGLGFSFVRITALLVAGSRLLVGTGNGVVISIPLETVWLHRRGQLGLRANKTSPTPS
GEGARPGCIITHVYGDSSDRRAASFTPYCSMAQALCFCHGRHDVAKFVSVPVNGVLAT
LNGSVLSDPAEGPGVLAAPASVEGQKRNVLVLSGEGYVDFRIGDGEDDDETEGAGD
MSQVQKPLVSKAFRSHLTWVQSVYTPR"

ORIGIN

Query Match	94.3%	Score 3482;	DB 9;	Length 5621;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 3489;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;
QY	180	GCCTCCCTGTCCTCGCAGCTGAGTCCCAACGGGGGCCAGGAGGACACCGGATGAAGAA	239	
DB	2124	GCCCGCCAAGTACAAAGCAGCTGAGTCCCAACGGGGGCCAGGAGACACCGGATGAAGAA	2183	
QY	240	CGTGCCCGGTGCCGGTGTACTGCCGCCCTCTGTGTGGAGAAGGACCCCAACCATGAAGCTGTG	299	
DB	2184	CGTGCCCGGTGCCGGTGTACTGCCGCCCTCTGTGTGGAGAAGGACCCCAACCATGAAGCTGTG	2243	
QY	300	GTGTGCCCGCGGCGTCTAACTGTAGCGGGTGGAGGCCCAATAGGAGACGACGCTGGGAATGG	359	
DB	2244	GTGTGCCCGCGGCGTCTAACTGTAGCGGGTGGAGGCCCAATAGGAGACGACGCTGGGAATGG	2303	
QY	360	AGTCAAGCCACGCGCAGCGCGGATCCCTGTGACTGCGACCGGCAAGGAGACGCGCAGGC	419	
DB	2304	AGTCAAGCCACGCGCAGCGCGGATCCCTGTGACTGCGACCGGCAAGGAGACGCGCAGGC	2363	
QY	420	CAAGAGCGCCACACGCTCTCCGAGAAAGAAAGGCCCAAGGAGCTCCCTGAAATGGACGC	479	
DB	2364	CAAGAGCGCCACACGCTCTCCGAGAGAGAGAGGCCCAAGGAGCTCCCTGAAATGGACGC	2423	
QY	480	CACCTCCAGCGGGTGTGGATCTCTGACAGACCCCTGACCAACAGCAAGGTGTGATCAT	539	
DB	2424	CACCTCCAGCGGGTGTGGATCTCTGACAGACCCCTGACCAACAGCAAGGTGTGATCAT	2483	
QY	540	CGAGCCCAACAGCCGGGCGACGGTGGTGGACAGTTCACGCTGTGCAACGCGACGCTGCT	599	
DB	2484	CGAGCCCAACAGCCGGGCGACGGTGGTGGACAGTTCACGCTGTGCAACGCGACGCTGCT	2543	
QY	600	GTGCATCTCCAGCATCCCGCGGCCACGCGACAGCGACTACCCCTCCCGGGGAGATGTTCCCT	659	
DB	2544	GTGCATCTCCAGCATCCCGCGGCCACGCGACAGCGACTACCCCTCCCGGGGAGATGTTCCCT	2603	
QY	660	GGACAGCAGCTGGAACCCAGAGGACCCCGGGCGCGAGTGGCGTGTGCGCGGTATCACCT	719	
DB	2604	GGACAGCAGCTGGAACCCAGAGGACCCCGGGCGCGAGTGGCGTGTGCGCGGTATCACCT	2663	

QY 1800 CGATGACAGCAGTGCAGGCGCCAGCAGCTTTCATCCCTACTGCTCCATGATGCCAGGC 1859
Db 3744 CGATGACAGCAGTGCAGGCGCCAGCAGCTTTCATCCCTACTGCTCCATGATGCCAGGC 3803
QY 1860 CCAGCTATGCTTCCATGCGGACCGCATGCGTGAAGTTCCTTGTCTCGGTGCCAGGGA 1919
Db 3804 CCAGCTATGCTTCCATGCGGACCGCATGCGTGAAGTTCCTTGTCTCGGTGCCAGGGA 3863
QY 1920 CGTGCTGGGCCACCTGAATGGCAGTGTGCTGGACAGCCAGCGAGGCGCTGGGCGAGC 1979
Db 3864 CGTGCTGGGCCACCTGAATGGCAGTGTGCTGGACAGCCAGCGAGGCGCTGGGCGAGC 3923
QY 1980 TGGCCCTGCTCGAGGTGCAGGCGCAGAAAGCTGCGAAAGCTGCTGCTGCTGAGCGGCGG 2039
Db 3924 TGGCCCTGCTCGAGGTGCAGGCGCAGAAAGCTGCGAAAGCTGCTGCTGCTGAGCGGCGG 3983
QY 2040 GGAGGGCTACATCGACTTCGCAATTGGACATGGACATGGACATGGACATGGACATGG 2099
Db 3984 GGAGGGCTACATCGACTTCGCAATTGGACATGGACATGGACATGGACATGGACATGG 4043
QY 2100 AGGGACATGAGCCAGGTGAAGCCCGTGTCTCCAAAGGAGAGCGCAGTCAATCATGCT 2159
Db 4044 AGGGACATGAGCCAGGTGAAGCCCGTGTCTCCAAAGGAGAGCGCAGTCAATCATGCT 4103
QY 2160 GTGGCAGGTGCTCTACACCCCGAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2219
Db 4104 GTGGCAGGTGCTCTACACCCCGAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4163
QY 2220 AGGACCCCGACCACTGACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 2279
Db 4164 AGGACCCCGACCACTGACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 4223
QY 2280 CCCTCTTTAACTCTCAACCTGACGTTTCACTGAGTGTGCGCCCTCCAGCGGCGAGG 2339
Db 4224 CCCTCTTTAACTCTCAACCTGACGTTTCACTGAGTGTGCGCCCTCCAGCGGCGAGG 4283
QY 2340 GAGTGGGGGATGCGGATCAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2399
Db 4284 GAGTGGGGGATGCGGATCAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4343
QY 2400 AGATGCTCTCGGACAGTTTCCGGGAGCTCTCGGAGAGTTCAGGCGCAGAGTCTCA 2459
Db 4344 AGATGCTCTCGGAGCAGTTTCCGGGAGCTCTCGGAGAGTTCAGGCGCAGAGTCTCA 4403
QY 2460 AGTCCAGGCGACCTTGGGCGCAGCGCAGCGAGCAATCCGAGGTGCTGCTGCTACCTG 2519
Db 4404 AGTCCAGGCGACCTTGGGCGCAGCGCAGCGAGCAATCCGAGGTGCTGCTGCTACCTG 4463
QY 2520 GGCCTCTACTCCCGCAGCACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2579
Db 4464 GGCCTCTACTCCCGCAGCACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4523
QY 2580 GCCCTGGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2639
Db 4524 GCCCTGGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4583
QY 2640 TGCCTAGGGAGCTGGGCGCAGGCACTAGCCTTTGCGCCAGGAGGTGGGCTCAGGCTGCC 2699
Db 4584 TGCCTAGGGAGCTGGGCGCAGGCACTAGCCTTTGCGCCAGGAGGTGGGCTCAGGCTGCC 4643
QY 2700 CAGTGTGCTGACCCAGCGGCTTCTGCTGGGCTTCCCGCTGCTCAAGCTCTATCCT 2759
Db 4644 CAGTGTGCTGACCCAGCGGCTTCTGCTGGGCTTCCCGCTGCTCAAGCTCTATCCT 4703
QY 2760 GTCTGTCCCGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2819
Db 4704 GTCTGTCCCGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4763
QY 2820 CTCCTTGGGCGAGCTGCTTGGAGACAGAGATGCTTACCCCATCTGCTGCTGCTGCTGCT 2879
Db 4764 CTCCTTGGGCGAGCTGCTTGGAGACAGAGATGCTTACCCCATCTGCTGCTGCTGCTGCT 4823
QY 2880 CTCCTTGGCAGCCCGCTTCTGACCCCGTGTCCCGCCAGGCTCTGCTGGGCGAGACTCAC 2939

Db 4824 CTCTTGCCAGCCCGTGTCTGACCCGCTGCCCGCAGGCTCTGCTGGGCAAGACTCAC 4883
QY 2940 CTTTGGAGGAGTGGGCGCTTGGAGTCTCTCCCTCCAGAGAGCCCGCCAGGCTTCTC 2999
Db 4884 CTTTGGAGGAGTGGGCGCTTGGAGTCTCTCCCTCCAGAGAGCCCGCCAGGCTTCTC 4943
QY 3000 AGGCTGCCAGGAGGAGGCGCTCAGGAAGAGGAGGAGGCGCTTGGCTCTCTCCGGAT 3059
Db 4944 AGGCTGCCAGGAGGAGGCGCTCAGGAAGAGGAGGAGGCGCTTGGCTCTCTCCGGAT 5003
QY 3060 CAGTCTTAGGACACAGGCTCAGCTCAGCTTGAATGGGAGTGAATGCTTCCCGGCGCTG 3119
Db 5004 CAGTCTTAGGACACAGGCTCAGCTCAGCTTGAATGGGAGTGAATGCTTCCCGGCGCTG 5063
QY 3120 CTTCTGACGCGGCTTCCACGAGGCGGCTTCCAGACACGCTTAAAGTGCCTAGGGTT 3179
Db 5064 CTTCTGACGCGGCTTCCACGAGGCGGCTTCCAGACACGCTTAAAGTGCCTAGGGTT 5123
QY 3180 GCGCGCTGTGGCTTGTCTCCAGGAGCAACAGAGGCGCAACAGAGGCGGCGGCTGGG 3239
Db 5124 GCGCGCTGTGGCTTGTCTCCAGGAGCAACAGAGGCGCAACAGAGGCGGCGGCTGGG 5183
QY 3240 CTGAGGATGAGCGCGCCCGCAGCGCTCCAGAGGCGCAACAGAGGCGGCGGCTGGG 3299
Db 5184 CTGAGGATGAGCGCGCGCCCGCAGCGCTCCAGAGGCGCAACAGAGGCGGCGGCTGGG 5243
QY 3300 CTGCTCTCCCTGCGGCTTGTAGGAGGCGGCTTGTAGGAGGCGGCGGCTGGGCTGGG 3359
Db 5244 CTGCTCTCCCTGCGGCTTGTAGGAGGCGGCTTGTAGGAGGCGGCGGCTGGGCTGGG 5303
QY 3360 TCTGCTGTGGAGGCGCTTGTAGGAGGCGGCTTGTAGGAGGCGGCGGCTGGGCTGGG 3419
Db 5304 TCTGCTGTGGAGGCGCTTGTAGGAGGCGGCTTGTAGGAGGCGGCGGCTGGGCTGGG 5363
QY 3420 AGCAGAGATCTTGGGAGGCTGCGCCACAGGCGGCTTGTAGGAGGCGGCGGCTGGG 3479
Db 5364 AGCAGAGATCTTGGGAGGCTGCGCCACAGGCGGCTTGTAGGAGGCGGCGGCTGGG 5423
QY 3480 GCGCTGCGGCTTCTCCCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3539
Db 5424 GCGCTGCGGCTTCTCCCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5483
QY 3540 GTCTGTCCCGACGATCTCGAAGTGAATCCGGGCTGAGCAGTGGGCGGCTGGGCGAGG 3599
Db 5484 GTCTGTCCCGACGATCTCGAAGTGAATCCGGGCTGAGCAGTGGGCGGCTGGGCGAGG 5543
QY 3600 GGTGACGATCTCTCTCAGGCTTGGCGCTTGCAGCAACCCACATATCTGCTCTGATGT 3659
Db 5544 GGTGACGATCTCTCTCAGGCTTGGCGCTTGCAGCAACCCACATATCTGCTCTGATGT 3659
QY 3660 AATAAATGCTTAAAGTC 3677
Db 5604 AATAAATGCTTAAAGTC 5621

RESULT 5
AB071076
LOCUS
DEFINITION Macaca fascicularis testis cDNA clone:QtsA-16678, full insert
AB071076
sequence.
AB071076
AB071076.1 GI:15451330
oligo capping; fis (full insert sequence).
KEYWORDS Macaca fascicularis (crab-eating macaque)
SOURCE Macaca fascicularis
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1
AUTHORS Osada, N., Hida, M., Kusuda, J., Tanuma, R., Hirata, M., Suto, Y.,
Hirai, M., Terao, K., Sugano, S. and Hashimoto, K.
TITLE Cynomolgus monkey testicular cDNAs for discovery of novel human

D	b		3824	AAGTTGAAGATGCACTGGTGTCTGAGTGGTGGTGAAGGTTACATTGACTTCCTCGTATCCGA	3883
Q	y		2068	GACGGAGAGGACGACGAGACGAGGAGAGGGCGCAGGGGACATAGCAGGTGAAGCCCGTG	2127
D	b		3884	GACGAGAGGATGATGAATAGAGAAATGTGCCGGGACGTGAACACAGAAAGCCCTCG	3943
Q	y		2128	CTGTCCAAGGCAGAGGAGTACATCATATCGTGTGGAGGTGTCTACACCCTGAGTG-	2187
D	b		3944	TGTCTCAAAGGCTGAGCGAGGCACATCATCGTGTGGCAGGTGTCTACACCCTGAGTG-	4002
Q	y		2188	AGCTCTCCCTGCCCTGCCAGACTGTACATAGGACCCCCCGACCACTGACCCCCTGCC	2247
D	b		4003	AGACCTGTCTTACCTGATGCCAATGTACATAGAACCTTACCTGCCTGCCCTCCCGCCT	4062
Q	y		2248	GGCCCGGGGTAGCCAGCCAGGCGCCCGCCCTCTTCTTAACCTCTCAACTGTCAGCT	2307
D	b		4063	GTTCCTGGGCGAGCA-----GGTGTCTCCATCCCTTTTAACTCTCAACTTTCAGCT	4117
Q	y		2308	TTACCTGAGTCTGCCCTCCAGCGGGCACGGAGTGGGGATGCGGATCAGCTGGGAG	2367
D	b		4118	TTTGCCTGAGGTCACGCCCCCTAGCTGTTAGAGAGGATGAAGCAGGTTTGGGGAGTGAG	4177
Q	y		2368	GAGGAGGGAGGGGTCTTCCACCGAGGGGAAATGCTCTCGGGAAGTTCCTCCGGGCA	2427
D	b		4178	CGAGAACAAGATGCTTTTCCCATCAGGGAGGAGAGCCCTCTGAGACAATCTCTCCAGCA	4237
Q	y		2428	GCTCTTGGCCAGCTTCCAGGCCACAGT-CCTCAAGTCCAGGGCACCTTGGGCCACGCGCA	2486
D	b		4238	GCCCTTGTCTTTGACCCNAACATGAAGTCCACAGGGGGCCCTCAGGGCAAGGACGACAAA	4297
Q	y		2487	GGCAGATCCAGTGTGCTCTTACCTTGGGCTCTCTAGCTGGGCTCTCTCCCGACGCTGGA	2546
D	b		4298	ATGACTATGGGCAGCCCCCTACAGGTGGCCCTGGCTCCACTCCATCTCACCACTCTGGC	4357
Q	y		2547	GGAGCAGGGGCTCCCCCGCCGAGGCTGCCTGCCTTGGGCCCACTCTGATGTG--	2604
D	b		4358	AGGAGCAGAGCCTCATACTGCCAGAGCTTCCCAAAGTCTACTTGTCTTTCATGAGCC	4417
Q	y		2605	CTCATGGGGCCACCTGCTTGGGCCCTCACTCTGCCTAGGGAGCTGGGCCACGCGAC	2664
D	b		4418	CTGAGGAAGGATCTCTGATTTCCAAACCTTTGTCTCTCTGAGGGACCCAGGCTTAGCCA	4477
Q	y		2665	TAGCCTTTGCCAGGAGGTGGGCTCAGGCTGCCAGGTSCCTGCAACCCAGCCCGGCTT	2724
D	b		4478	CCATCTCTGCCAGAAACAAAGTTAGGGCTACCTGGTTCCCGGGGCCCGACGACCT	4537
Q	y		2725	TCTCTGGGGCTCCCGCTGCTCAAGCTCTATCTGTCTGTCTCCCAACCCAGCTGTCCCG	2784
D	b		4538	TCTCT- GGGCTTCCCGCTCTCAAACCTTACATTTCCCTTGTCTCCCTAGCTGGGTCTTAC	4596
Q	y		2785	TGCCAGGAGCTGGCCATAAAGCACAGAGCCCGGCTCCCTGGGGCAGCTGCTTGAGAAC	2844
D	b		4597	TG---GGAACCTGGCATGAAGCTCAGGGCCAGCTCCCAAGGGCAGCTGCTT---CAAA	4650
Q	y		2845	AGAGCTGTACCCCATCTGCCCATGACGGCAGGCTTTGCCAGCCCGCTTCTGACCCG	2904
D	b		4651	AGATGACAAATAGCCGTCAACCCAAAGATTGTCCAGCCCTTCAAGACCAAGTT-----	4702
Q	y		2905	TGTCCCCCAGGCTCTGCCTGGGCGAAGACTCACTTGGAGGAGTGGGCCCTTGGAGTCC	2964
D	b		4703	----CCCTCGGCTCTTCTTGGGACAAAGATTACCTTGGAGGAG-----CAAGTGCC	4752
Q	y		2965	TGTCCCTCCCAAGCCCCCAGAGTGGGATTTCTCAGGCTGCCAGGGCAGGCCAGCT	3024
D	b		4753	TGCCCTCCCAAGACCCCCAAGGTTAGATTTCTCAGGCTGCCAGGGCAGGCACAGGCT	4812
Q	y		3025	CAGGAAGAGGGAGGCCCTTGGCCTCTCCGGGATCAGTCTTAGGACACAGGCTCAGCT	3084
D	b		4813	CAGGAAGAGGG--TGSCCTTGTCTCCCAAGAGTCAGTCTCTTAGATGAGGCTTAGCT	4870
Q	y		3085	CAGTTGATGGGGATGATGTGCTCCCGGGGCTGCTCTCTGCAAGGGGCTCCAGGAGC	3144

Db	4871	CAGGTTAATCGAGGATGGTGTGCTCCCAAGGCGCTGCCTCTCGACAGGGCTCCAGGAGC	4931
Qy	3145	CCAGCTCCACAGACAGCTACTAAGTGCCTAGGAGTTCGCCCGCTGTGGCTCTCCACAGGG- 3203	
Db	4931	CCAGCTCCACAGACAG- TACTAAGTGCCTAGGAGTTCGCCCGCTGTGGCTCTCCCTTTGGA 4989	
Qy	3204	----AGCAACAGAGAGGGCCACCAAGCAGAGGCCGCTGGGGCTCAGGATCGAGCGGCCCC 3259	
Db	4990	GAGCAGCTGACAAGAGGGCTCCAGCAGAGGCCAGAAAGCTGAGGA-----CACGCC 5042	
Qy	3260	AGCCGACTCCAAAGCCCGCAGAGGSCACAGCCACCCCTGGACTGCTCTCCCTGCGCCAGCTG 3319	
Db	5043	AGGATGGGCACAGTCTTAAGCAAGTGTGTGGAGCAGGTGCTGCCCTGGCTCTGGCCCTTG 5102	
Qy	3320	GGCCTCTCTGGCCTATTCTTACCTTCAGGCCCACTGCATCTCCTGTCTGGAGGCCCTTGA 3379	
Db	5103	CTCGTCAGGGGC-----A 5115	
Qy	3380	TGAGGGCAGCCCGACCCCGCACCCACCCCAACCCAGAGAGCACAGATCTTTGGGAGCT 3439	
Db	5116	CTGAGTCAGTTGGCTCTACTGCAGGTTCTTACCAGAGAACACAGATCCAGGGGCACT 5175	
Qy	3440	GCCCCACAAAGCCCGCTGGCCACCGAGGGCTGCAGC--CGCTCGCTGCGGGCTTCTCCC 3497	
Db	5176	GCCCC-TGAGGCTGTCTGGCCACAAAGCTGCAGCAGTGTGCAGCTGCAGGCTTCTCCC 5234	
Qy	3498	CACCACCTGGCACCTCCTCATGTGATGTATGTCCGCTTCCTCTGTGTTCCTCCAGGATC 3557	
Db	5235	CACCACCTGGCACCTCTCCACTGTGATGTATG---TCCCTGTCTGTTCCTCCACATGAT 5291	
Qy	3558	TCGAAGTGACTCCGGGCTGAGCAGTGGGGCGGCTGGGGAGGGGTG-----ACGATTC 3611	
Db	5292	GTGAGCTGACAGGGTTAGTTGGAGGGTGCAACTGGAGGAGGGGCCAGTGCAACTTT 5351	
Qy	3612	CCTCAGGCTTTGGCCCTGCAAGCAAAACCCACATATCTGCTCTGTATGTAATAATGCTTT 3671	
Db	5352	CCTCAGGCTTTGGCTGCTGCAAGTGAACCTCACATATCTGCTCTGTATGTAATAATGCTTT 5411	
Qy	3672	AACGCTGTAATAAAAAAAAAAAAAA 3694	
Db	5412	AACACTAATAAAAAAAAAAAAAA 5434	
RESULT 7			
AF178636			
LOCUS	AF178636	5443 bp	linear
DEFINITION	Mus musculus JNK interacting protein-3a (Jip3)	linear	ROD 21-JAN-2000
ACCESSION	AF178636		
VERSION	AF178636.1		
KEYWORDS	GI:6724091		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 5443)		
TITLE	Kelkar,N., Gupta,S., Dickens,M. and Davis,R.J.		
JOURNAL	Interaction of a mitogen-activated protein kinase signaling module with the neuronal protein JIP3		
MEDLINE	Mol. Cell. Biol. 20 (3), 1030-1043 (2000)		
PUBMED	20094982		
REFERENCE	10629060		
AUTHORS	2 (bases 1 to 5443)		
TITLE	Kelkar,N., Gupta,S., Dickens,M. and Davis,R.J.		
JOURNAL	Direct Submission		
FEATURES	Submitted (17-AUG-1999) Molecular Medicine, HHMI & UMMS Medical School, 373 Plantation Street, Worcester, MA 01605, USA		
source	Location/Qualifiers		
	1. .5443		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57/B6"		
	/db_xref="taxon:10090"		
	/issue_two="three"		

```
gene 1. 5443
CDS /gene="Jip3"
41. 4009 /gene="Jip3"
/function="JNK Map kinase signal transduction pathway
molecular scaffold"
/notes="JIP3a; putatively alternatively spliced"
/codon_start=1
/product="JNK interacting protein-3a"
/protein_id="AAF26842.1"
/db_xref="GI:6724092"
/translation="MMETOMDEGGGVVYVDDYCGSVMSERVSLAGSIYREPERLI
HCYDEEVVKELMPLIVNLENSVLENOHEVELELLREDNQQLTOYERERKALRI
QAEKFIIEFDALQSEKKBELQIQVEHYEFQRLKAKKNYADQISRLERESBEMKE
YNALQHRTEMIVTEVHEIERSKMQVGGSGQTESLSPGRSKRKPSTLNVFPLADGM
VRAQMGLKVPAGDHLSDLGQLQSSSYQCNDMESESGSAAATPSTGTGKSN
PTSPVSAATVPLNLESLOPLGVYVTKNKKAREKNRNVEQVOTEMRNELVIMGM
SSDWSVDQDIIDSTPELDVCPETRLERTGSPQGIYNKAFINTSLYHLSHTAGS
EVI GDVDEGADLLGEFGMGKEVGNLLLENSQLLETKNALNVVKNDLIAKVDLSGEQ
EVLGELEAAKQAKVLENIKELEBELKRVKSEAVTARREPREVEDDKIPMAORRE
FTRVEMARVLMNQIKERLMELQEAVRTEMIRASREHPSVQEKKASTIWOQFSRLIF
SSSSPPPAKSYRNVNTHYKSPNTAGFSQRRSHALCOISAGSRPLEFFDDDDCTSSA
RREKREQYRVRHVRNDGRLOACGWSLPKQKQLSPNGSQEDTRMKNVPVYPCR
PLVEKDPSTKLWCAGNLSGWKPHEDSSNGPKVPQGRDPLTCDREGEPEPKSTHPS
PEKKAKETPEADATSSRVMLTSLTTSKVILIDANOPGTIVDQTVCNHVLICSS
IPAASDSIYPGEMFLDSVNPEDSGADVLGILTVGCATRONVPRNSCSRGDTPV
LDKGGDVATANGKVNQSPTEATEATEVDPGPSPSEATVTRPGELTEHVFTDPA
PTPSSSTQAPASENGSNTIVQPOVPEPSGELSTTSAAPTMILGAQNGWLYVHSVA
ANWKILSIKDSVLSLVHVKGRVLVALDGLTALFHRGEDGQWDLNHYHMDLGH
PHHSIRCMANVDRWCYKKNVHVIOPTQWIOIEKSFDAHPRRSQVRLAWIGDGVW
VSI RLDSTRLYHATHOHLQVDIEPVSKMLGTGKLGFSFVRITALLIAGNRLVWG
TGNGVSIPIPTETVTLHRGOLLGIRANKTSPTSGEGTRPGSIHIVYGDSDSKAAS
FITPYSMAQAGCFHGRHDVAKFPVSVFQNVILATLNGSVLSPSESGPAPADAEG
QKLNALVLSGEGYIDPRIGDEDETEECAGDVNQTPLSKAERSHIIVWQSVYT
PE"
```

Query Match 48.7%; Score 1797.4; DB 10; Length 5443;
Best Local Similarity 72.9%; Pred. No. 1.1e-253; Indels 108; Gaps 18;
Matches 2599; Conservative 0; Mismatches 856;

QY	148	GATGGGTAGAGCCAGGGTTCCTGCGCCACGGCGCTCCCTGCTCCCTGCGAGCTGAGTCCTC	207
DB	1970	GACGGAGGCTGACGCGCTGTGGTGGAGCTGCGCTGCGCAAGTACAGACGAGCTGAGCCCC	2029
QY	208	AACGGGGCCAGAGGACACGGGATGAAGAAGCGTGCCTGGTGCCTGCTGAGCTGAGTCCTC	267
DB	2030	AATGAGGCGCAGAGACACCGGATGAAATAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2089
QY	268	CTGGTGGAGAGACCCACCATGAGCTGTGTGCGCGGGGCTCAACCTGAGCGGG	327
DB	2090	CTGGTGGAGAGACCCCTTCGACAAAGCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2149
QY	328	TGGAGGCCCAATGAGGACGACGCTGGGAATGGAGTCAAGCCAGCGCCGAGCCCGGATCCC	387
DB	2150	TGGAAGCCACATGAGAGGACTCTAGCAATGAGCCCAAGCCCTGTACCAAGCTCGAGACCT	2209
QY	388	CTGACCTGCGACCGGAGAGGACGCGGAGCCCAAGAGCGGCCACACGCTCTCCGAGAG	447
DB	2210	CTGACCTGTGACCGGAGAGGAGAGGCGAACCACCAAGAGCACACACCATCACCTGAGAAG	2269
QY	448	AAGAGGCCCAAGAGCTCCCTGAAATGGAGCGCACCTCCAGCGGGGTGTGGATCTCTGACC	507
DB	2270	AAGAAGCAAGAAACCCCTGAGGAGATGTACTCTCAGTGGGTATGGATCTCTCACC	2329
QY	508	AGCACCTTGACCAACAGAGAGGTGTGTGATCATCGAGCCCAACAGCGGGCGACGTTGGTG	567
DB	2330	AGCACCTTGACCAACAGAGAGGTGTGTGATCATCATGATGATGATGATGATGATGATGATGAT	2389
QY	568	GACCACTTACCGTGTGACAGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	627
DB	2390	GATCAGTTACAGTGTGCAATGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2449

ORIGIN

QY	628	GACAGCAGTACCCTCCGGGGAGATGTTCTTGACAGCGAGCGTGAACCCAGAGAGACCG	687
DB	2450	GACAGTGAATATCCCCCTGGGGAGATGTTCTTACAGAGTGAATGTAACCCCTGAAGATTCA	2509
QY	688	GGCCAGATGCGTGTGGCGGGTATCACCTGTGTGGGTGTGACACCCGCTGCAACGTG	747
DB	2510	GGTCTGATGTTGTGTGGTGGCATCACCTGTGTGGGTGTGTGTACCCGCTGCAATGTT	2569
QY	748	CCGGGAGCACTGCTCTCTCCGGGGGACACCCAGTGTGTAGACAAGGGGCGACGGGGAG	807
DB	2570	CCAGTGAACACTGTTCTCTACAGGAGACACCCAGTGTGTAGACAAGGGGCGACGGGGAT	2629
QY	808	GTGGCCACCATGCGCAACGGGAAGGTCAACCCGCTCCAGTCCACAGAGAGGCGCCACAG	867
DB	2630	GTGGGACCATGCGCAATGGGAAGGTCAACCCGCTCCAGTCCACAGAGAGGCGCCACAG	2689
QY	868	GCCACGGAGGTGCGAGACCCCTGGGGCCAGCGAGCCAGAGACGACACATTGGGGCCGGG	927
DB	2690	GCCACAGAGGTGCGAGACCCCTGGGGCCAGCGAGGTCAAGAGCAACGACAGTCCGGGCCGGG	2749
QY	928	CCTCTCAGAGCAGTCTTCTCACTGACCCCGACCCCGCTCTCTTGGGCCCCAGCCT	987
DB	2750	CCTCTCAGAGCAGTCTTCTTACTGACCCAGACCCCGCTCTCTTGGGCCCCAGCCT	2809
QY	988	GCGACGAGAAACGGGCCAGAGCCTGACAGCAGCAGCACAACGCGCCAGAGCCAGAGCCACG	1047
DB	2810	GCCAGTGAATGGTGAATCCAAATGGCACTTGTACAGCCTCAGGTGGAGCCAGT	2869
QY	1048	GGGACCCCAACGGGAGGAGGAGCAGTCTGACCCACCATGTGCTGGAGCCCGAGAAC	1107
DB	2870	GGGGAACCTCTCAACAAACACAGTAGCCTGCACCCACTATGTGGCTAGGAGCCAGAA	2929
QY	1108	GGTGGCTCTATGTCACTCGCTGTGCGCAACTGGAAGAGTGTGCTGCTCAATCAAG	1167
DB	2930	GGCTGGCTCTATGTGCACTCAGGTAGCAGCACTGGAAGAGTGTGCTGCTCAATCAAG	2989
QY	1168	CTGAAGGATCTGTGCTGAGCCTGGTGCATGTCAAAGCGGTGTGCTGCTGCTGCTGCTGCTG	1227
DB	2990	CTAAGAGACTCTGTGCTGAGCCTGGTGCATGTCAAAGCGGAGTGTGCTGCTGCTGCTGCTG	3049
QY	1228	GACGGGACCTTGGCCCATCTTCCACCGTGTGAGAGTGGCCAGTGGAGTGTGAGCACTAT	1287
DB	3050	GATGGGACCTTGGCTATCTTCCATCGTGGAGAGTGGCCAGTGGAGCCTGAGCAACTAC	3109
QY	1288	CACCTAATGGACTCTGGGCCACCCGACCACTCCATCCGCTGCATGGCTGTGTGTGACGAC	1347
DB	3110	CACCTAATGGACTCTGGGCCACCCGACCACTCCATCCGCTGCATGGCTGTGTGTGATGAC	3169
QY	1348	CGCGTGTGTGGCTACAAGAACAAAGGTGCACGTATCCAGCCCAAGACCATGCAATA	1407
DB	3170	CGAGTTTGTGTGGCTACAAGAACAAAGGTGCATGTTATCCAGCCCAAGACCATGCAATA	3229
QY	1408	GAGAAGTCAATTTGAGCCCAACCCGCGGGGAGAGCCAGGTGGGAGCTGGCGTGGATC	1467
DB	3230	GAGAATCAATTTGATGCCCCACCCAAAGGGGGAGAACAGGTACGTGAGTGGCTGGATC	3289
QY	1468	GGCGATGGGTATGGGTGTCCATCCGCTGGAATCCACCTCAGGCTCTTACCATGCACAC	1527
DB	3290	GGTGTGAGTGTGGTGTCTTATTGCTGTGATTCTACCTTCGGCTCTTACCATGCTCAC	3349
QY	1528	ACGCAACAGCATCTACAGGACGTGACATTTAGAGCCCTAGTCAAGCAAGTGTAGGCACT	1587
DB	3350	ACCCACACAGCATCTGAGGATGTGACATTTAGAGCCCTATGTTAGCAAGATGCTAGGAAC	3409
QY	1588	GGCAAGCTGGGTTCTCTTCTGTAGCATCAGGCCCTGCTTGTGCGGGGAGCCGGCTC	1647
DB	3410	GGCAAGCTGGGTTCTCTTCTGTAGCATCAGGCCCTGCTTGTGCGGGGAGCCGGCTC	3469
QY	1648	TGGGTGGGCAACGGCAACCGAGTGTGATCTCCATCCCCCTGACAGAGACTGTGTGCTCTG	1707
DB	3470	TGGGTGGGCACTGGCAATGGGGTGTGATCTCCATCCCCCTGACAGAGACTGTGTGCTCTG	3529
QY	1708	CACCGAGGCCAGTCTCTCTGGGGCTCGAGCCAAATAGACATCCCCCACTCTGGGGAGGGC	1767

Db	3530	 CATGAGCCAGCTCTCTAGGCTCCGAGCCAAAGACATCCCCAAATCTGGGAGGGG
QY	1768	 GCCCGTCCCGGGGCATCATCCAGTGTATGGCATGACAGCAGTACAGGGGGGCCAGC
Db	3590	 ACCGCCAGGGGCAATCATCAATGTGTATGGGACACAGCATGAAAGGCCGCCAGT
QY	1828	 AGCTTCATCCCTACTGTCTCCATGGCCGACAGCCGAGCTATGCTTCCATGGCACCCGAT
Db	3650	 AGTTTCATCCCTACTGTCTCATGGCAACAGCTCAGCTTGTCTTCCATGGCACCCGAT
QY	1888	 GCCGTGAAGTTCTTGTCTCGGTCCAGGGAACGTGTGGCCACCTCTGAATGGCAGTGTG
Db	3710	 GCTGTCAATTTCTTTGTCTCTGTGCCAGGAATGTCTGTGCCACTCTCTCAATGCGCAGTGTG
QY	1948	 CTGACAGCCGACGCGAGGGCCCTGGGCGACGCTCCCTCGAGGTCTCGAGGGCCAG
Db	3770	 CTACAGAGCCCATCAGAGGGCCCTGGGCTGCTGCAACCGCTGAGATGTGAGGGCCAG
QY	2008	 AAGCTGCGGAACGTGTGTGTCTGAGCGGGGAGGGCTACATCGACTTCCCATTTGGA
Db	3830	 AAGTTGAAGAAATGCACTGTGTGTGAGTGTGTGAGAGTTACATGTACTTCCGTATCGGA
QY	2068	 GACGAGAGACACAGACGAGAGGGGGCGGAGGACATGACCCAGTGTGAAGCCCGTG
Db	3890	 GACGAGAGATGATGAATCTGAGGAATGTGCGGGGACGTGAACCAAGAGCCCTCG
QY	2128	 CTGTCCAGGCGAGCGCAGCTCAATCATCGTGTGGCAGGTGTCTTACACCCCGAGTGA
Db	3950	 TTGTCCAAAGCTGAGCGCAGCCACATCATCGTGTGGCAGGTGTCTTACACCCCTGAGTG-
QY	2188	 AGCTGTGCTGCTGCTGCTGCTGCTGTAATAGAACCCCGACCACTGACACCCCGCC
Db	4009	 AGACCTGTCTTACTCTGATGCCAACTGTATATAGGACCTTACCTGCTGCTCCCGCCT
QY	2248	 GGCCCGGGGTAGCCAGCCAGCGCGCGCGCTCTCTTACCTCTCAACCTGACAGCT
Db	4069	 GTTCCCTTGGGGACGCA-----GGTTCGTCCATCCCTCTTTTAACTCTCAACTTGCAGCT
QY	2308	 TTCACTTGAATGTGGCCCTCCAGCGGCGAGGAGTGGGGGATGGGATCAGCTGGGAG
Db	4124	 TTTGCTGTAGGTCAGGCCCTAGCTGTGTAGAGAGGATGAGACAGGTTTGGGAGTGAG
QY	2368	 GAGGAGGGAGGGTGTCTTCCACCGAGGGGAAGATGCTCTCGGACAGTTCCTCCGGCA
Db	4184	 ACAGGAACAGATGTCTTTCCATCAGGGAGAGAGAGCTCTGAGACAACTCTCCAGCA
QY	2428	 GCTCTCTGGCCAGCTTCCAGCCACAGT-CCTCAAGTCCAGGACACCTTGGGCCAGCGCA
Db	4244	 GCCCTTGTCTTACCCACATGAATCTCCAGGGGCCCTCAGGGCAAGCGAGACCAAA
QY	2487	 GGCAGAAATCGAGGTGTCTGTGCTTACCTTGGGCTCTCTATCTCCCCAGCAACCTGGA
Db	4304	 ATGACTATGGGACAGCCCTACAGGTGGCTGGCTGCTTCCCTTCCACTCTCACCAGTCTTGGC
QY	2547	 GGAGGAGGGGCTCCCGCGCGAGCTGCTGCTGCTGGGCCCACTCTGATGTG--
Db	4364	 AGGACAAGAGCTCATCTGCCAGAGTCTTCCCAAGTCTACTTGTCTTGTGATGCAGCC
QY	2605	 CTCATGGGGCAACCTGCTCTGGGCCCTCACTCTGCTTAGGGAGCTGGGCGAGGCAC
Db	4424	 CTAGAGAGGGCTCTTGATTTCCAAACCCCTTGTCTCTCTGAGGGGCCAGAGCTTAGCCA
QY	2665	 TAGCCCTTGGCCAGGAGGTGGGCCCTCAGGCTGCCAGGTGCTGACCCCAAGCCGGCT
Db	4484	 CCGTCTCTGCCAGAAGCAAGTTAGGGGTACCTGTGTTCCGGGGCCCCAGGACGCT
QY	2725	 TCTCTGGGCGCTCCCGTGTGAAGCTCTATCTCTGTGTCTCCCAACCCCACTGTCC
Db	4544	 TCTCT--GGGCTTCCCGCTCTCAAACTTATATCCCTTGGCCCTTAGCCTGGGTCTTACC
QY	2785	 TGCCAGGAGCTGGCAATAAGCACAGGCCCGGCTCCCTGGGAGAGCTGCTTGAAGAC

Db	4603	TG-----GGAACCTGGCATGAAAGCTCAGGCGCAGCCTCCCAAGGGCAGCTGCTTCAAAG	4655
QY	2845	AGAGACTGCTACCCATCTCTGCCCATCGAGCGAGGCTTGTGCACGCCCCCTTTGTGACCCCG	2904
Db	4659	A-----TGACAAATACCCGTCACACCAAGATTGTCCAGGCCCTTCACAGACCAG	4706
QY	2905	TGTCCTCCCAAGGCTCTGCTGGGCAGAAAGTCACTTCTGGAGAGATGGGCCCTGGAGTCC	2964
Db	4707	TTCCCTCC--GCTTCTCTGGACACAAGATTCACTTTGGAGAG-----CAAGTGCCT	4758
QY	2965	TGTCCTCTCCCAAGACCCCAAGGTGGATTTCTCAGGCTGCCAGGCGCAGGCCACAGGCCT	3024
Db	4759	TGCCCTCTCCAGAAAGCCCAAGGTGAATTTCTCAGGCTGCCAGGCGCAGGCACAGGCCT	4818
QY	3025	CAGAAAGAGGGAGGCCCTTGGCCTCTCCGGGATCAGTCTCTAGGACACAGGCTCAGCCT	3084
Db	4819	CAGAAAGAGG--TGCCCTCTGGTCTCCAGAGTCACTCTAGGATGCAAGGCTTAGCCT	4876
QY	3085	CAGTTTATGGGGATGATGTCTCCCGGGCTGCTCTCTGACCGGGCTCCACAGGAGC	3144
Db	4877	CAGTTTATGGAGATGGTGTCTCCAGAGGCTGCTCTCTGACAGGCTCCAGAGGCT	4936
QY	3145	CCAGCTCCCAAGACCGTACTAAAGTCTAGGTTTCCCGCTGTGGCCTGTGCCCTGCTCCAGG-	3203
Db	4937	CCAGCTCCCAAGACAG--TACTAAGTGCTAGGTTTCCCGCTGTGGCCTGTGCCCTTGA	4995
QY	3204	----AGCAAAGAGAGGCCACCAAGCAGAGGCCCTGTGGGCTGAGATGAGCGCGCCCT	3259
Db	4996	GAGCAGCTGACAAGAGGCTTCCAGGAGAGGCCAGAGAGCTGAGGA-----CCAGCC	5048
QY	3260	AGCCGACTCCAAGCCCGCAGAGGGCAGAGCCACCTGGACTGCTCTCTCCCTGCCAGCTG	3319
Db	5049	AGGATGGGCACAGTTCTAAGCAAGTGTGTGGAGCAGGTGCTGCCCTGGCTCTGGCCTTG	5108
QY	3320	GGCTCTCTGGCCTATTCCTTACCTTCAGGCCCACTGCACTCCTGTCTGGAGGCCCTTA	3379
Db	5109	CTGCTCAGGGCC-----A	5121
QY	3380	TGAGGGCAGCCAGCCCGCACCACCCACCAACCAAGAGAGCAGATCTTGGGAGGT	3439
Db	5122	CTGAGTCAGTTGGCTCTACTGCAACCTTCCCAAGAGAGCAGATCCAGGGGCACT	5181
QY	3440	GCCCCCAAGCCCGCTGGCCACCGAGGCTGCAGC--CGTGGCTGCGCGCTTCTCC	3497
Db	5182	GCCCC--TGAGCTGTCTGGCCCAACAAGCTGCAGATGTCAGCTGAGGCTTCTCC	5240
QY	3498	CACACCTGCGACCTCCACTGTGATGTATGTCCGCTCCCTGCTGTGTTCGCCCAAGATC	3557
Db	5241	CACACCCCTGCCACTCTCCACTGTGATGTATG--TCCCTGTCTGTTCCACATGAGT	5297
QY	3558	TCCAAGTGATCCGGGCTGAGCAGTGGGGGCTGGGGAGGGGTG-----ACGATCT	3611
Db	5298	GTGAGTGACCGGGTTAGTTGGAGGGGTGCAACTGGAGGGAGGGGCGCAGTGCAACTTT	5357
QY	3612	CCTCAGGCTTTGGCCCTGCAAGCAACCCACATATCTGCTGTATGTAAATAAATGTCTT	3671
Db	5358	CCTCAGGCTTTGGCCCTTGAAGTGAACTCACATATCTGCTGTATGTAAATAAATGTCTT	5417
QY	3672	AACGTGTAATAAAAAAAAAAAAAA	3694
Db	5418	AACACTAAAAAAAAAAAAAAAAAAAA	5440
RESULT 8			
AB093282			
LOCUS			
DEFINITION			
AB093282			
ACCESSION			
VERSION			
AB093282.2			
KEYWORDS			
FLI_CDNA.			
SOURCE			
Mus musculus (house mouse)			
ORGANISM			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AB093282			
Mus musculus mRNA for mKIAA1066 protein.			
AB093282			
Mus musculus			
5411 bp			
mRNA			
linear			
ROD 15-FEB-2003			

QY	1288	CACCTAATGGACCTGGGCCACCGCACCACTCCATCCGCTGCATGCGCTGTGTGTACGAC	1347
DB	3097	CACCTAATGGACCTGGGCCACCGCACCACTCCATCCGCTGCATGCGCTGTGTGTAAATGAC	3156
QY	1348	CGCGTGGGTGTGGCTACAAGAACAAAGGTGCACGTCATCCAGGCCAACAGACCATGCAGATA	1407
DB	3157	CGAGTTTGGTGTGGCTACAAGAACAAAGGTGCATGTTATCCAGCCCAAGACAAATGCAGATT	3216
QY	1408	GAGAAATCATTTGACGCCCAACCGCGGGGAGAGACAGGTGCGGCGAGCTGGCGTGGATC	1467
DB	3217	GAGAAATCATTTGATGCCCCACCAAGCGGGAAAGCCAGGTACGTCAGCTGGCCTCGATC	3276
QY	1468	GGCGATGGCGTATGGGTGCCATCCGCTGGACTCCACCTTGAGGCTCTACCAATGCACAC	1527
DB	3277	GGTATGAGTGTGGTCTTAATTCGTTGGATTCTACCTTGGCTCTACCATGCTCTAC	3336
QY	1528	ACGCACAGCATCTACAGGACGTGGACATTTAGGCGCTTACGTACGAAGATGCTAGGCACCT	1587
DB	3337	ACCACCAAGCACCTCGAGGATGTGGACATTTAGGCGCTTATGTTAGCAAGATGCTAGGAACC	3396
QY	1588	GGCAAGCTGGGTTCCTCCTTGTAAGCATCAAGCGCCCTGCTTGTGCGGGGAGCGCGGCTC	1647
DB	3397	GGCAAGCTGGGCTTCCTCCTGTCGGCATCACAGCCTTACTCATGCGAGCAACCGTCTG	3456
QY	1648	TGGTGGGCACCGGCAACGGAGTGGTTCATCTCCATCCCTCGACAGAGACTGTGCTCCTG	1707
DB	3457	TGGTGGGCACCTGGCAATGGGGTGTTCATCTCCATCCCTTGACTGAGACTGTGCTCTG	3516
QY	1708	CACCGAGGCGAGCTCTGGGGCTCCGAGGCCAATAAGACATCCCCCACTCTGGGGAGGGC	1767
DB	3517	CATCGAGGCGAGCTCTTAGGGCTCCGAGGCCAACAAAGACATCCCCAACTCTGGGGAGGGG	3576
QY	1768	GCCCGTCCCGGGGCAATCATCCAGTATATGGGATGACAGCAGTGACAGGGCGGCCAGC	1827
DB	3577	ACCCGCCAGGGGGCATCATCATGTGTATGGGAGCAGCAGTGCACAGGCGCGCCAGT	3636
QY	1828	AGCTTCATCCCTACTGCTCATGGCCAGCGCCAGCTATGCTTCCATGGGCACCCCGAT	1887
DB	3637	AGTTTCATCCCTACTGCTCCATGGCACAGGCTCAGCTTTCCTTCCATGGGCACCGTGAT	3696
QY	1888	GGCGTGAAGTTCTTTGCTCGGTGCCAGGAAAGTGTGTCGCCCACTTGAATGSCAGTGTG	1947
DB	3697	GCTGTCAAAATCTTTGCTCTGTGCCAGGAAATGTGCTGGCCCACTCTCAATGGCAGTGTG	3756
QY	1948	CTGACAGGCCAGCGAGGGCCCTGGGCCAGCTGCCCTCGAGGTGCGAGGCCAG	2007
DB	3757	CTAGACAGCCCATAGAGGGCCCTGGGCCCTGTGSCACCCGCTGCAGATGCTGAGGGCCAG	3816
QY	2008	AAGTCGGGAACGTGCTGTGCTGAGCGCGGGGAGGGCTACATCCACTTCGCATTGGA	2067
DB	3817	AAGTTGAAGAAATGCACTGGTGTCTGAGTGTGTGTAAGTTACATTTGACTTCCGTATCGGA	3876
QY	2068	GACGGAGAGACAGCAGAGCGGAGGAGGGCGCAGGGGACATGAGCCAGGTGAAGCCCGTG	2127
DB	3877	GACGGAGAGGATGATGAAATGTAGGAAATGTGCGGGGAGCTGAAACAGACAAAGCCCTCG	3936
QY	2128	CTGTCCAAGGACAGCGAGTGCATCATCTGTGTGGCAGGTGTCTTACACCCCGAGTGA	2187
DB	3937	TTGTCCAAGGCTAGGCGCAGCACAATCATGTGTGGCAGGTGTCTTACACCCCTGAGTG-	3995
QY	2188	AGTGTGCCCCCTGCGGCCGACCTGTATAGGACCCCCGACACCTGACCCCCGCCC	2247
DB	3996	AGACCCCTGTCTTACCTGATGCCAACTGTATAGGACCCCTACCTGCTGCTCCCTCCCGCT	4055
QY	2248	GGCCCGGGGTAGCAGCAGCGCGGCCCTCTTCTTAACCTCTCAAACCTGAGCT	2307
DB	4056	GTTGCCCTGG---GGCAGCGCAGGTGTGTCATCCC-CTTTTAACTCTCAAATGTCAGCT	4111
QY	2308	TTCACTGAGTCTGGCCCCCTCCAGCGGCGAGGAGTGGGGATGCGGATCAGCTGAGGAG	2367
DB	4112	TTTGCTGAGGTCACGCCCTTAGCTTTAGAGAGGATGAAGCAGGGTTTGGGGAGTGA	4171

QY	2368	GAGGAGGGAGGGGTGCTTCCACCCGAGGGGAAGATGCTCTCGGGACAGATTGTTCCCGGGCA	2421
DB	4172	GCAGGAACAGATGTTCTTTCCCATCAGGGAGGAGAGAGCTCTGAGACAAATCTCTCCAGCA	4231
QY	2428	GCTCTGTGCCAGCTTCCAGGCCAGAGT-CCTCAAGTCCAGGGACCTTGGGCCCAGGCCA	2486
DB	4232	GCCCTGTCTTTGACCCAAATGAAAGTCCACAGGGGGCCCTCAGGGCAAGCAGGACCAAA	4291
QY	2487	GSCAGAATCCGAGGTGTCTTGGCTCTACCTCTGGGCTCTACTCTCCCGACGACCCCTGGGA	2546
DB	4292	ATGACTATGGCAGCCCTCAGAGTGGCCCTTGGCTCACTCCATCTCACCAGTCTCTGGC	4351
QY	2547	GGAGCAGGGGCTCCCGCCCGCAGAGGTGCTGCTCCCTGGGCCCACTCTGCATGTG--	2604
DB	4352	AGGAGCAAGAGCCCTACTGTCCAGAGTCTTCCCAAGTCTACTGTCTTTGCATGCAGCC	4411
QY	2605	CTCATGGGGCCACCTGCTCTGGGCCCTCACTCTGCTTAGGGAGCTGGGCCCAGGCAC	2664
DB	4412	CTGAGGAAGGCTCCTGATTTCCAACTTTGTCTCTGTAGGGGGCCAGGCTTAGCCA	4471
QY	2665	TAGCCTTTGCCACGAGGAGTGGGCTCAGGCTGCCAGGTGCCTGCAACCCAGCCCGGCTT	2724
DB	4472	CCGTCTCTGCCACAGAAGCAAGTTTAGGGCTACCTGTGTTCCCGGGGCCCCAGGACGCT	4531
QY	2725	TCTCTGGGCTCCCGCTGTCGAAGCTCTATCTGTCTGTCTCCCAACCCCAAGCTGTGCC	2784
DB	4532	TCTCT-GGGCTTCCCGCTCTCCAACTTAATTTCCCTTGGCCCTTAGCCTGTGGTCTTACC	4590
QY	2785	TGCCACGGAGCTGGCATAAAGCACAGAGCCCGGCTCCCTGGGGCAGCTGCTTGAGAAC	2844
DB	4591	TG-----GGAACCTTGGCATGAAGCTCAGGGCCAGCTCCCAAGGGCAGCTGCTTCAAAAG	4646
QY	2845	AGAGACTGCTACCCCATCTCGCCCATCGACGCAAGGCTCTTGCCAGGCCCGCTTCTGACCCG	2904
DB	4647	ATG-----ACAAATACCGGTCACCCAAAGATTGTCCAGGCCCTCTCAGACCAAG	4694
QY	2905	TGTCCTCCAGGCTCTGCTGGGCGAAGACTCACTTGGAGAGTGGGCCCTTGGAGTCC	2964
DB	4695	TTCCCTCC--GCTCTCTCTGACACAAGATTCACTTGGAGG--CAAGTGCC--	4746
QY	2965	TGTCCTCCAGAGCCCCAGAGTGGGATTTCTCAGGCTGCGCAGGGCAGGCCAGGCT	3024
DB	4747	TGCCCTTCCAGAGCCCAAGGGTGAATTTCTCAGGCTGCCAGGGCAGGCACAGGCT	4806
QY	3025	CAGGAAGAAGGGAGGCCCTTGCCCTCTCGGGATCAGTCTTAGGACACAGGCTCAGGCT	3084
DB	4807	CAGGAAGAAGG--TGSCCTGTCTCCCAGAGTCACTCTAGATGACAGGCTTAGCT	4864
QY	3085	CAGTTGATGGGGATGATGTGCTCCCGGGCTGCTCTCTGACGGGCTCCACGGAGC	3144
DB	4865	CAGTTAATGGAGATGGTGTGCTCAAGGGCTGCTCTGCAAGGGCTCCAAGGAGC	4924
QY	3145	CCAGTCCCAGACACGCTACTAAGTGCCTTAGGTTGCCCGCTGTGGCTGTCTCCAGG-	3203
DB	4925	CCAGTCCCAGACAG--TACTAAGTGCCTTAGGTTGCCCGCTGTGGCTGTCTCCCTTGA	4983
QY	3204	-----AGCAACAGAGAGGCCCAAGCAGAGGCCGTGGGGCTTGAGATGGAGCCGCCGCC	3259
DB	4984	GAGCAGTGCACAGAGGCTTCCAGGCAGAGGCCAGAGAGCTGAGA-----CCAGCC	5036
QY	3260	AGCCGACTCCAGCCCGCAGAGGCGAGCCCACTCTGATGTCTCTCCCTGCGCAGCTG	3319
DB	5037	AGGATGGGCACAGTCTAAGAGCAAGTGTGTGGAGCAGGTTGTCCTTGGCTCTGGCTTG	5096
QY	3320	GGCTCTCTGGCTATTCTCTACCTTCCAGGCCCATGCACTCTCTGTCTGGGAGGCCCTTA	3379
DB	5097	CTCGTCAAGGGC-----A	5109
QY	3380	TGAGGGCAGCCAGCCCGCCAGCCCAACCCCAACAGAGAGCAAGATCTTGGGAGCT	3439
DB	5110	CTGAGTCAGTTGGCTCTACTGACGCTTCCACAGAGAGACAGATCAGGGGCACT	5169
QY	3440	GCCCAACAAGCCCGCTGGCCACAGAGGCTGACG--CGTGGGCTGCGGCTTCTCCC	3497

QY 988 GGCAGCGAGACGGGCGAGAGCCTGTGACAGCAGCAGCACAGGCCAGCCAGCCAGC 1047
Db 2931 GCCAGTGAGATGGGTGAGATCCATGGCAATGTTGACAGCTCAGGTGAGGCCAGT 2990
QY 1048 GGGGACCCACAGGAGCAGGAGCAGTGTGACACCACTATGCTGGGAGCCAGAAC 1107
Db 2991 GGGGAACCTCAACAACAACAGTAGTGGCTGCACCCACTATGTTGGCTAGGAGCCAGAA 3050
QY 1108 GCGTGGCTATGTGACCTCGGTGGCCACTGGAAGAGGCTGCTGGTGGCTGGCG 1167
Db 3051 GCGTGGCTATGTGATTCAGCGGTAGCCACTGGAAGAGTGTCTGACCTCCATCAAG 3110
QY 1168 CTGAAGGATTCGTGCTGAGCCTGTGCTATGCAAAAGGCGGTGTGCTGGTGGCTGGCG 1227
Db 3111 CTAAAGAGACTGTGCTGAGCCTGTGCTATGCTGATGCAAAAGGCGGTGTGCTGGTGGCTGGCG 3170
QY 1228 GACGGACCCCTGGCCATCTTCCACGTGTGTAAGATGGCCAGTGGGATCTGAGCAACTAT 1287
Db 3171 GATGGACCCCTGGCTATCTTCCATCGTGGAGAGGATGGCCAGTGGGACCTGAGCAACTAC 3230
QY 1288 CACCTAATGGACCTGGGCCACCGCAGCACCTCCATCCGCTGATGGCTGTGTTGTACGAC 1347
Db 3231 CACCTAATGGACCTGGGCCACCGCAGCACCTCCATCCGCTGATGGCTGTGTTGTACGAC 1347
QY 1348 GCGGTGTGGTGTGCTACAGAACAGGTGCAAGTGTGCACTCCAGCCCAAGACCAATGCAGATA 1407
Db 3291 CGAGTTGGTGTGGCTACAGAACAGGTGCAATGTTATCCAGCCCAAGCAATGCAGATT 3350
QY 1408 GAGAGTCAATTGACGCCACCGCGGGGGAGAGCGCAGGTGGCGAGTGGCGTGGATC 1467
Db 3351 GAGAAATCATTTGATGCCCCACCCAAAGGGGGAAGCCAGGTACGTGAGTGGCTGGATC 3410
QY 1468 GCGATGCGTATGGGTGTCCATCCGCTGGACTCCACCTGAGGCTTACCATGCAAC 1527
Db 3411 GGTGATGAGTGGGTCTTATTCGCTTGAATTCACCTTCGGCTTACCATGCTCAC 3470
QY 1528 AGCACAAGATCTACAGGACGTGGACATGAGCCCTAGCTCAGCAAGATGCTAGGCAC 1587
Db 3471 ACCCAGCAGCCTCGAGGATGTGACATGAGCCCTATGTTAGCAAGATGCTAGGAAC 3530
QY 1588 GGCAGCTGGTTCCTCCTGTTAGCGATCAGGCCCTGCTGTGCGGGCAGCGGCTC 1647
Db 3531 GGCAGCTGGGCTTCCTCCTGTTGGCATCAGCCCTTACTCATGCAAGCAACCGTCTG 3590
QY 1648 TGGGTGGGACCGGCAAGGAGTGTGATCTCCATCCGCTGACAGACACTGTGTCCTG 1707
Db 3591 TGGGTGGGCACTGGCAATGGGGTGTCTCTCCATCCGCTTGAAGTGTGACTGTGTCCTG 3650
QY 1708 CACGAGGCCAGCTCCTGGGCTCCGAGCCAAATAGACATCCCCCCTCTGGGAGGGC 1767
Db 3651 CATCGAGGCCAGCTCCTAGGCTCCGAGCCAAAGACATCCCCCAATCTGGGAGGGG 3710
QY 1768 GCCGTCGGGGGATCATCCAGTGTATGGGATGACAGATGACAGGGGCGGACG 1827
Db 3711 ACCGCGCCAGGGGATCATCCATGTGTATGGGAGCAGCAGTGTGCAAGGCGCCAGT 3770
QY 1828 AGCTTCATCCCTACTGTCTCAGCCAGCCAGCTATGCTTCCATGGGACAGCGGAT 1887
Db 3771 AGTTTCATCCCTACTGTCTCAGCCAGCCAGCTATGCTTCCATGGGACAGCGGAT 3830
QY 1888 GCGGTGAAGTTCCTTGTCTGGTGCAGGGAACTGTGTCGACACCTGAAATGGCAGTGTG 1947
Db 3831 GCTGTCAAAATCTTGTCTGTGTCAGGAAATGTGTGCGCACTCTCAATGGCAGTGTG 3890
QY 1948 CTGACAGCCAGCGAGGCGCTGGCCAGTGGCCAGTGGCCAGTGGGAGGTGAGGGCCAG 2007
Db 3891 CTAGACAGCCCATCAGAGGGGCTGGGCTGCTGCAAGGCTGAGATGCTGAGGGCCAG 3950
QY 2008 AAGCTGGGAACGTGGTGTCTCAGCGGGGAGGGCTACATCGACTTCCGATTTGGA 2067
Db 3951 AAGTTGAAGATGCACTGGTGTCTGAGTGGTGGTGAAGGTTACATTTGCTTCCGATTCGGA 4010

QY 2068 GACGGAGGAGCAGCAGAGCAGGAGGAGGCGCAGGGGACATGAGCCAGGTGAAGCCCGTG 2127
Db 4011 GACGGAGAGGATGATGAAACTGAGGAATGTGCGGGGACGTGAACACAGACAAGGCCCTCG 4070
QY 2128 CTGTCTCAAGCAGAGCGCAGTCACTCATCTGTTGTGCGAGTGTCTTACACCCCGAGTGA 2187
Db 4071 TTGTCCAGGCTGAGGCGAGCCACATCATCTGTTGTGCGAGTGTCTTACACCCCTGAGTG- 4129
QY 2188 AGCTGTGCTGCTGCTGCGCCCGACCTGTATAGGACCCCGACCACTGTAGACCCCGCC 2247
Db 4130 AGACCTGTCTCTACCTGATGCACTGTATATAGGACCCCTACTGCTGCTGCCCTCCCGCCT 4189
QY 2248 GSCCGCGGGTAGCGCAGCCAGCGCGCGCCGCTCTTCTTAACCTCTCAACCTGAGCT 2307
Db 4190 GTTCCTGTGGGAGCA-----GGTCTGCTCATCCCTTTTAACTCTCAACTGTGAGCT 4244
QY 2308 TTCACTGTAGTGTGCTGCTGCTGCGGGGATGCGGGGATGCGGATCAGCTGCGGAG 2367
Db 4245 TTTCCTGTAGGTGTCAGCCCTTAGCTGTTAGAGAGGATGAACAGGGTTTGGGAGTGAG 4304
QY 2368 GAGGAGGAGGGGTGCTTCCACCCGAGGGGAGATGCTCTCGGACAGTCTTCCCGGCA 2427
Db 4305 GCAGGAACAGATGTCTTCCCATCAGGGAGGAGAGAGCTCTGAGACATCTCTCCAGCA 4364
QY 2428 GCTCTGCGCAGCTTCCAGCCCGCAGAGT--CCTCAAGTCCAGGCACTTGGGCGCCAGCGCA 2486
Db 4365 GCGCTGTGTCTTGAACCAACATGAAGTCCAGGGGCGCTCAGGGCAAGGAGGAGCAACAA 4424
QY 2487 GGCAGAAATCCAGGTGTCTGCTGCTTACCTTGGGCTCTTACTTCCCGCAGCACCCCTGGA 2546
Db 4425 ATGACTATGGGAGCGCCCTCAGGTGGCGCTGGCTCCACTCCATCTCACCAGTCTTGCG 4484
QY 2547 GGAGCAGGGGTCCCGCGCGAGGCTGCTGCTGGGCGCCACCTCTGATGCTG-- 2604
Db 4485 AGGAGCAGAGAGCTCATCTGCCAGAGTCTTCCCAAGTCTACTTGTCTTGCATCAGCG 4544
QY 2605 CTGATGGGCGCACCTGCTGCTGGCGCTCAGTCTGCTAGGAGGAGTGGGCGCAGGAC 2664
Db 4545 CTGAGGAAGGCTCTGATTTCCAACTTATATCCCTTGGCGCTAGCTGGTCTTACC 4604
QY 2665 TAGCCTTTGCCAGGAGGTGGGCTCAGGCTGCGCAGTGCCTGACCCAGCGCGCT 2724
Db 4605 CCGTCTCTGCGCAGAGCAAGTTAGGGCTACCTGTTCCCGGGGCGCCAGGAGCT 4664
QY 2725 TCTCTGGGCGCTCCCGCTGCTCAAGCTCTATCTGCTGTCTGCTCCACCCAGCTGCTCCC 2784
Db 4665 TCTCT--GGGCTTCCCGCTCTCAAACTTATATCCCTTGGCGCTAGCTGGTCTTACC 4723
QY 2785 TGCCCGAGGAGCTGGCATAAAGCAGAGGCGCGGCTCCCTGGGCGCAGCTGCTTGAGAAC 2844
Db 4724 TG-----GGAACCTGGCATGAAAGCTCAGGGGCGAGCTCCCAAGGCGAGCTGCTTCAAAAG 4779
QY 2845 AGAGACTGTACCCCATCTGCGCCATGCGAGCAGGCTCTTCCAGCGCCCGTCTCTGACCCG 2904
Db 4780 ATG-----ACAAATACCGCTCAACCAAGATTGCTCAGCGCTCAGAGCCAG 4827
QY 2905 TGTCCCGCAGGCTCTGCTGGGAGAGACTCACTTGGAGGAGTGGGCGCTGAGTCC 2964
Db 4828 TTCCCTCC--GCTCTTCTTGGACACAAGATTACCTTGGAGGAG-----CAAGTGCC 4879
QY 2965 TGTCCCTCCAGAGACCCCGAGGGTGGATTCTCAGGCTGCCAGGCGCAGGCCAGCGCT 3024
Db 4880 TGCCCTTCCAGAGACCCCGAGGGTAGAATTTCTCAGGCTGCCAGGCGCAGGCCAGCGCT 4939
QY 3025 CAGGAAGAGGGAGGCGCTGCTGCTCGGATCAGTCTTAGGACACAGGCTCAGCT 3084
Db 4940 CAGGAAGAGGG--TGCGCTGTCTCTCCAGAGTCTCTAGATGAGGCTTAGCT 4997
QY 3085 CAGGTTGATGGGAGTATGCTCCCGGCGCTGCTCTGACCGGGCTCCACGAGC 3144
Db 4998 CAGGTTAATGGAGATGGTGTGCTCCAAAGGCTGCTGCTCTGACAGGGCTCCAGAGGAGC 5057
QY 3145 CCAGCTCCCGACAGCAGCTACTAAGTGCCTAGGGTGTGCCCGCTGTGGCTGCTCTCCAGGG- 3203

|||||
5058 CCAGTCCCGACACG-TACTAAGTGCCTAGGCTTGCGGCTGTGGCTGCTCCTCTTGGGA 5116
QY
3204 ----AGCAACAGAGAGGCGCACCAAGACAGAGCCCGTGGGCTGAGGATGAGAGCGCGCCCC 3259
Db GAGCAGCTGACAAAGAGGCTTCAGGCAGAGGCGCAGAGAGCTGAGGA-----CCAGCC 5169
QY AGCCGACTCCAAAGCCCGCAGAGGCGCAGAGCCACCTGAGATGTCTCTCCCTGCGCAGGTG 3319
Db AGGATGGGCACAGTCTTAAGCAAGTGTGTGGAGCAGGTGCTGCCCTGGCTCTGGCCTTG 5229
QY GGCTCTCTGGCCTATTCCTACCTTCCAGGCCACTGCACTCCTCTCTGGGAGGCCCTTA 3379
Db CTCGTCAGGGGC-----A 5242
QY TGAGGGCAGCCAGCGCCCGCACCCACCCACCAACAGAGAGCAGATCTTGGGGAGCT 3439
Db CTGAGTCAGTTGGCTCTACTGACGCTTCCACAGAGAGCAGATCCAGGGGCACT 5302
QY GCGCCACAGCCCGCTGGCCACCGAGGGCTGCAGC--CGCTGCGCTCGCGCTTCTCCC 3497
Db GCGCC--TCAGGCTGTCTGCGCCACAAAGCTGCAGAGTGTGCAGCTGCAGGCTTCTCCC 5361
QY CACCACTCTGCACCTCCACTGTGATGTATGTCCTCCCTGCTGTTCCCGCAGGATC 3557
Db CACCACTCTGCACCTCTCCACTGTGATGTATGTCCTCCCTGCTGTTCCCGCAGGATC 5418
QY TCGAAGTGACCTCGCGGCTGAGCAGTGGGCGCGCTCGGGGAGGGGTG-----ACGATTCT 3611
Db GTGAGTGCACGGGGTGTAGTTGGAGGGTGCACCTGGAGGAAGGGCGCAGTGCACATT 5478
QY CCTCAGGCTTGGGCGCTCAAGCAACCCACATATCTGCTCTGATGT 3659
Db CCTCAGGCTTGGGCGCTCAAGTGAACCTACATATCTGCTCTGATCT 5526

RESULT 10
HS371H6
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
HS371H6 20612 bp DNA linear PRI 19-APR-2001
Human DNA sequence from clone LA16-371H6 on chromosome 16 Contains
part of the gene for KIAA0516 protein, part of a gene similar to
NDP kinase, ESTs, an STS and 2 CpG Islands, complete sequence.
AL031718.11 GI:13937341
HTG; CpG Island; KIAA0516; kinase.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 20612)
Baggaley, C.
Direct Submission
Submitted (07-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On May 3, 2001 this sequence version replaced gi:13277292.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/LA16-371H6 is

part of a clone contig from the tip of the short arm of chromosome
16 spanning 2Mb of pl3.3 (Higgs D.R., Flint J., Daniels R., MRC
Molecular Haematology Unit, Institute of Molecular Medicine, Oxford
(unpublished)), and is from the Los Alamos, flow sorted human
chromosome 16 libraries constructed by Norman Doggett
(unpublished). IMPORTANT: This sequence is not the entire insert of
clone LA16-371H6 It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.
The true left end of clone LA16-371H6 is at 1 in this sequence. The
true left end of clone LA16-447E6 is at 20513 in this sequence. The
true right end of clone LA16-361A3 is at 3304 in this sequence.
FEATURES
Location/Qualifiers
1..20612
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="LA16-371H6"
/clone_lib="LA16"
join(269..451,529..718,838..957,1032..1112,1213..1280,
1458..1699,1873..2044,2114..2227,2378..2416,2508..2687,
2786..2935,3012..3133)
/gene="c371H6.1"
join(<269..451,529..718,838..957,1032..1112,1213..1280,
1458..1699,1873..2044,2114..2227,2378..2416,2508..2687,
2786..2935,3012..3133)
/gene="c371H6.1"
/note="match: CDNAs: Em:AB011088
match: ESTs: Em:AI815491 Em:AI816436 Em:AA984532
Em:AA013330
match: proteins: Tr:O60271 Wp:CE00363 Tr:O15013"
/codon_start=3
/evidence=not_experimental
/product="c371H6.1 (KIAA0516 protein)"
/protein_id="CAB72318.1"
/db_xref="GI:6941067"
/db_xref="GOA:Q9UPT6"
/db_xref="SWISS-PROT:Q9UPT6"
/translation="ASDSDFYPPGEMFLDSVNPEDPGADGVLGAGTILYGCATRCNVPR
SNCSRGTPTVLDKGGEVATANGKVPNSQTEATEATEVDFGPGSEPTATLRPG
PLTFHTFTDPAPTPSSGPGSGSENGPEPSSTPEPEPSGDTGAGSAPATMLWGA
QNGLYVHSVAANWKIKLSIKLSDSVLSLVHVKRVLVALADGTALAI.FHRGDEGQWL
LSNVLMDLGHPHHSIRCMVAVYDVRCYKKNVHIQPKTMQIEASAGQGGEGKRL
LLASSPASSIRKSTPKYKSDAHPRESOYROLAWIGDGVVWSIRLSDTLRLYHAHT
HQHQVDIEFVSKMLGTGKLGFSFVITALLVAGSLRWGTGNGVLSIPLTETVV
LHRGQLGLRANKTSPTSGEGAPGDIHVYDSDSDRAASFPFYCSMAQQLCFHG
HRDAVKFVSPGNVLATNGSLVDSPAEGPAPAPASEVEGQKLRNLVLVGGEGYI
DFRIGDGEDDETEGAGDMSQVKPVLKAERSHIIVQVSYTPE"
4620..8266
/note="CpG island"
/evidence=not_experimental
complement(join(4958..5075,5190..5302,5379..5480,
5589..5719))
/gene="c371H6.2"
complement(join(4958..5075,5190..5302,5379..5480,
5589..5719))
/gene="c371H6.2"
/note="match: CDNAs: Em:LI6785 Em:S39901 Em:U61287
Em:AF033377 Em:U89605 Em:M36981 Em:U85511 Em:M91597
Em:X92956 Em:X92957 Em:U29656 Em:X97900 Em:X97901
Em:X68193 Em:X58965 Em:AF043543 Em:M55331 Em:AF043542
Em:M35970 Em:AF045187 Em:M65037 Em:X97899
match: ESTs: Em:AA999847 Em:W11654 Em:AA387534 Em:W39079
Em:W48246 Em:W83979 Em:AI525854 Em:AI609754 Em:HB3355
Em:AA276203 Em:W83990 Em:AI356825 Em:AI799023 Em:AA129462
Em:AI173169 Em:AA781191 Em:AI587608 Em:AA399289
match: proteins: Sw:P15532 Sw:P15531 Sw:P19804 Sw:Q90380
Sw:P52174 Sw:P52175 Tr:O57535 Sw:Q05982 Sw:Q13232
Sw:P27950 Tr:P70071 Sw:P26550"
/codon_start=3
/evidence=not_experimental
/product="c371H6.2 (similar to NDP kinase)"


```
/protein_id="CAB72319.1"
/db_xref="GI:6941068"
/db_xref="GOA:Q3NUP9"
/db_xref="SPTREMBL:Q3NUP9"
/translation="CTGAHERITFLAVKPDGVORRLVGEIVRRFERKFKLVALKLVAQ
SEELREHYAEILREFFYGRKLVKYMAGPVPVAMVWGLDVRTSRALIGATNPADAPP
GTRIGDFCIEVGEKNLTHGSDSVESARREIALWFRDELLCWEDSAGHWLYE"
5536..5589
/note="27 copies 2 mer cc 70% conserved"
6068..6125
/note="29 copies 2 mer cc 69% conserved"
13032..13462
/note="TIM3e repeat: matches 426..872 of consensus"
complement(15564..16020)
/note="match: STS: Em:G37311"
16033..18297
/note="CpG island"
/evidence=not_experimental
16198..16313
/note="58 copies 2 mer cc 59% conserved"

ORIGIN
Query Match 47.4%; Score 1751.4; DB 9; Length 20612;
Best Local Similarity 91.7%; Pred. No. 3.8e-247;
Matches 1946; Conservative 0; Mismatches 1; Indels 174; Gaps 2;

QY 1734 AGCCAAATAAGACATCCCCACCTCTGGGAGGCGCCCGTCCCGGGGGCATCATCCAGT 1793
Db 2506 AGCCAAATAAGACATCCCCACCTCTGGGAGGCGCCCGTCCCGGGGGCATCATCCAGT 2565
QY 1794 GTATGGCGATGACAGAGTGAAGAGGCGGCCAGCAGCTTTCATCCCTACTGCTCCATGGC 1853
Db 2566 GTATGGCGATGACAGAGTGAAGAGGCGGCCAGCAGCTTTCATCCCTACTGCTCCATGGC 2625
QY 1854 CCAGGCCAGCTATGCTTCCATGGGACCGCGATGCCGTGAAGTCTTTGTCTCGGTG--- 1911
Db 2626 CCAGGCCAGCTATGCTTCCATGGGACCGCGATGCCGTGAAGTCTTTGTCTCGGTGTC 2685
QY 1912 ----- 1911
Db 2686 AGGTGAGGCTGGGCCCCCTCTGCCATCCACATCCCTCTGCATGCCAGTGGCGCGCCCTCC 2745
QY 1912 ----- CCAGGAAAGCTGTGTCGCCACCCCTG 1935
Db 2746 CCCAGAGGCGCTGTCTGTAATCGCTTCTGCCATCCAGGAAAGCTGTGTCGCCACCCCTG 2805
QY 1936 AATGCACTGTGCTGGACAGCCAGCCAGGCGCCCTGGGCGAGCTGCCCTGCTCGGAG 1995
Db 2806 AATGCACTGTGCTGGACAGCCAGCCAGGCGCCCTGGGCGAGCTGCCCTGCTCGGAG 2865
QY 1996 GTCAGGGCCAGAGAGCTGCGGAAAGCTGTGCTGTGCTGAGCGGCGGAGGCTACATCGAC 2055
Db 2866 GTCAGGGCCAGAGAGCTGCGGAAAGCTGTGCTGTGCTGAGCGGCGGAGGCTACATCGAC 2925
QY 2056 TTCCGCAAT----- 2064
Db 2926 TTCCGCAATGCTGAGCGGCGCCAGGAGACAGGCTGAGGTTGGGCGGGGGAGCCCTGG 2985
QY 2065 ----- GGAGACGAGAGACGACGAGAGCGGAGCGC 2099
Db 2986 CCCTCACTCTGCTTTGCGCCGAGGAGACGAGAGGACGACGAGAGCGGAGGCGC 3045
QY 2100 AGGGACATAGCCAGGTGAAGCCGCTGCTTCCAAAGCCAGAGCGCAGTCAATCATCGT 2159
Db 3046 AGGGACATAGCCAGGTGAAGCCGCTGCTTCCAAAGCCAGAGCGCAGTCAATCATCGT 3105
QY 2160 GTGCAAGGTGTCTTACACCCCGAGTGAAGTGTGCTGCCCTGGCCCGCCAGCTGTACAT 2219
Db 3106 GTGCAAGGTGTCTTACACCCCGAGTGAAGTGTGCTGCCCTGGCCCGCCAGCTGTACAT 3165
QY 2220 AGGACCCCGACACTGACCCCGCGCGCGCGGGGTAGCCAGCCAGCGCGCGCGC 2279
Db 3166 AGGACCCCGACACTGACCCCGCGCGCGGGGTAGCCAGCCAGCGCGCGCGC 3225
```

2280 CCCTTTCTAACTCTCAACCTGCACTTTCACCTGAGTGTGCGCCCTCCAGCGGGCAGG 2339

3226 CCCTTTCTAACTCTCAACCTGCACTTTCACCTGAGTGTGCGCCCTCCAGCGGGCAGG 3285

2340 GAGTGGGGGATGCGGATCAGCTGGAGGAGGAGGAGGGGTGCTTCCACCCGAGGGGA 2399

3286 GAGTGGGGGATGCGGATCAGCTGGAGGAGGAGGAGGGGTGCTTCCACCCGAGGGGA 3345

2400 AGATGCTCTCGGGACAGTTCCTCCGGGCAGCTCTCTGCGCAGCTTCCAGGCCAGAGTCTCA 2459

3346 AGATGCTCTCGGGACAGTTCCTCCGGGCAGCTCTCTGCGCAGCTTCCAGGCCAGAGTCTCA 3405

2460 AGTCCAGGCGACCTTGGGCCCGCAGCGCAGGACAAATCCGAGGTGGTCTCTGCTCTACCTG 2519

3406 AGTCCAGGCGACCTTGGGCCCGCAGCGCAGGACAAATCCGAGGTGGTCTCTGCTCTACCTG 3465

2520 GGCCCTCTACTCCCGCAGACCCCTGGAGGAGGAGGGGTCCCGCGCGCGAGGTGCTCT 2579

3466 GGCCCTCTACTCCCGCAGACCCCTGGAGGAGGAGGGGTCCCGCGCGCGAGGTGCTCT 3525

2580 GGCCCTGGGCCACCTCTCTGATGCTCTATGGGGCCACCCCTGCTCTCTGGGCCCTCACTC 2639

3526 GGCCCTGGGCCACCTCTGATGCTCTATGGGGCCACCCCTGCTCTCTGGGGCCCTCACTC 3585

2640 TGCCTAGGGAGTGGGCCAGGCACTAGCTTTGTCCTGCGAGGAGTGGGCTCAGGCTGCC 2699

3586 TGCCTAGGGAGTGGGCCAGGCACTAGCTTTGTCCTGCGAGGAGTGGGCTCAGGCTGCC 3645

2700 CAGGTGCTGCAACCCAGCGCGCTTCTGCGGGCTCCCGTCTCAAGCTCTATCTT 2759

3646 CAGGTGCTGCAACCCAGCGCGCTTCTGCGGGCTCCCGTCTCAAGCTCTATCTT 2759

2760 GTCTGTCCCGACCCAGCTGTCTGCGCGAGTGGCATAAAGCAGGAGCCCGG 2819

3706 GTCTGTCCCGACCCAGCTGTCTGCGCGAGTGGCATAAAGCAGGAGCCCGG 3765

2820 CTCCTGGGGAGTGTGTTGAGAAACAGAGTGTCTACCCCATCTCTGCGCCATCAGGCGAG 2879

3766 CTCCTGGGGAGTGTGTTGAGAAACAGAGTGTCTACCCCATCTCTGCGCCATCAGGCGAG 3825

2880 CTCCTGGCGAGCCGCTGCTGACCCGCTGCTCCCGCAGGCTCTGCTGGGCGAGAGTCTAC 2939

3826 CTCCTGGCGAGCCGCTGCTGACCCGCTGCTCCCGCAGGCTCTGCTGGGCGAGAGTCTAC 3885

2940 CTTGGAGAGTGGGCCCTGAGTCTCTCTCTCCAGAGAGCCCGCAGGCTGGGATTTCTC 2999

3886 CTTGGAGAGTGGGCCCTGAGTCTCTCTCTCCAGAGAGCCCGCAGGCTGGGATTTCTC 3945

3000 AGGCTGCCAGGCGAGGCCCTCAGGAAAGAGGGAGGCCCTCTGCGCTCTCCGGGAT 3059

3946 AGGCTGCCAGGCGAGGCCCTCAGGAAAGAGGGAGGCCCTCTGCGCTCTCCGGGAT 4005

3060 CAGTCTTAGGACACAGCTCAGCTCAGTGTGATGGGGATGATGCTCTCCGGGGCTG 3119

4006 CAGTCTTAGGACACAGCTCAGCTCAGTGTGATGGGGATGATGCTCTCCGGGGCTG 4065

3120 CCTCTTGCACGGGGCTCCACGGAGCCAGCTCCAGACACAGCTACTAAGTCCCTAGGGTT 3179

4066 CCTCTTGCACGGGGCTCCACGGAGCCAGCTCCAGACACAGCTACTAAGTCCCTAGGGTT 4125

3180 GCCCGTGTGGCTGTCTCCAGGAGCAACAGAGAGGCCAACAAAGCAGAGGCCCGGTGGG 3239

4126 GCCCGTGTGGCTGTCTCCAGGAGCAACAGAGAGGCCAACAAAGCAGAGGCCCGGTGGG 4185

3240 CTGAGGATGAGCGGCCCGCAGCGACTCCAGCCCGCAGAGGCGACAGCCACCTTGA 3299

4186 CTGAGGATGAGCGGCCCGCAGCGACTCCAGCCCGCAGAGGCGACAGCCACCTTGA 4245

3300 CTGCTCTCTCTGCGCGCTCTCTGCGCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3359

4246 CTGCTCTCTCTGCGCGCTCTCTGCGCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4305

Qy 3360 TCCTGTCTGGAGGCGCTTATGAGGGCAGCCAGCCCGCCCGCACCACCCCAACACGAGA 3419
 Db 4306 TCCTGTCTGGAGGCGCTTATGAGGGCAGCCCGCCCGCACCACCCCAACACGAGA 4365
 Qy 3420 AGCAGAGATCTTGGGAGCTCCCAAGCCCGCTGGCCACCGAGGGCTGCAGCCGCT 3479
 Db 4366 AGCAGAGATCTTGGGAGCTCCCAAGCCCGCTGGCCACCGAGGGCTGCAGCCGCT 4425
 Qy 3480 GCGTGTCCGGCTTCTCCCAACACCCCTGCCACCTCCACCTGATGATGATGTCGCTCCCTC 3539
 Db 4426 GCGTGTCCGGCTTCTCCCAACACCCCTGCCACCTCCACCTGATGATGATGTCGCTCCCTC 4485
 Qy 3540 GTCTGTCTCCCAAGATCTCGAAGTACTCGGGCTGACGATGGGGGGCTGGGGGAGG 3599
 Db 4486 GTCTGTCTCCCAAGATCTCGAAGTACTCGGGCTGACGATGGGGGGCTGGGGGAGG 4545
 Qy 3600 GGTGACGATCTCTCTCAGCTTTTGCCCTGCAAGCAAAACCCACATATCTGCTCTGATGT 3659
 Db 4546 GGTGACGATCTCTCTCAGCTTTTGCCCTGCAAGCAAAACCCACATATCTGCTCTGATGT 4605
 Qy 3660 AATAAATCTTAACGTCGTA 3680
 Db 4606 AATAAATCTTAACGTCGTA 4626

RESULT 11
 AE006639 270150 bp DNA linear PRI 15-AUG-2002
 LOCUS Homo sapiens 16p13.3 sequence section 7 of 8.

DEFINITION Homo sapiens 16p13.3 sequence section 7 of 8.
 ACCESSION AE006639 AE005175
 VERSION AE006639.1 GI:14336758

KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 270150)
 Daniels,R.J., Peden,J.F., Lloyd,C., Horsley,S.W., Clark,K.,
 Tufarelli,C., Kearney,L., Buckle,V.J., Doggett,N.A., Flint,J. and
 Higgs,D.R.

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

Sequence, structure and pathology of the fully annotated terminal 2
 Mb of the short arm of human chromosome 16
 Hum. Mol. Genet. 10 (4), 339-352 (2001)
 21096910
 11157797
 2 (bases 1 to 270150)
 Daniels,R.J., Peden,J.F., Lloyd,C., Horsley,S.W., Clark,K.,
 Tufarelli,C., Kearney,L., Buckle,V.J., Doggett,N.A., Flint,J. and
 Higgs,D.R.

TITLE
 JOURNAL
 Institute of Molecular Medicine, John Radcliffe Hospital, Oxford,
 Oxon OX3 9DS, UK

FEATURES
 Location/Qualifiers

1..270150
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16p13.3"

1..5613
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="cosmid 425c2"
 1..37670
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="cosmid 315g5"
 53799..93949
 /organism="Homo sapiens"
 /mol_type="genomic DNA"

source
 source
 source

source

source

source

misc_feature

repeat_region

repeat_region

misc_feature

gene

CDS

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

gene

mRNA

/db_xref="taxon:9606"
 /clone="cosmid 431h6"
 89167..121683
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="cosmid 329f2"
 113820..160576
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="cosmid 361a3"
 157079..193613
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="cosmid 371h6"
 <1..25277
 /note="HS395F10"
 complement(54..172)
 /note="FRAM; RepeatMasker predicted
 23/10/2000"
 /rpt_family="SINE/Alu"
 190..477
 /note="AluSg; RepeatMasker predicted
 23/10/2000"
 /rpt_family="SINE/Alu"
 1371..4511
 /note="CpG island; K1AA0590"
 /evidence=not_experimental
 complement(1479..4124)
 /gene="gsl14"
 complement(join(1479..1729,1981..2170,3687..4124))
 /gene="gsl14"
 /note="GENSCAN prediction; no supporting evidence known at
 this time"
 /codon_start=1
 /product="unknown"
 /protein_id="AAK61287.1"
 /db_xref="GI:14336759"
 /translation="MGTKPRPEASCGCTGCGSDAPPALSRARQPPAGPRHDS
 AAQWSPFSTPGSLASRGLPGCQNGRAAPAPRRLLQGRNAGLVVRIGWS
 GASPTSPRAGARLGRITQILPPLTAKPAPRQQQFLRRPSGITWPTLGRSC
 TCVPTDLRLFFAARGASVLAASLTFGYFRVPGAPLSQRAAQRGRAGRPLEWA
 GGGRRHTPEAKDKREMQNMCLVTRWVGEGRVASNRQRMASQYRPLDAAT
 LRKXALDGLGEGAL"
 2542..2836
 /note="AluDb; RepeatMasker predicted
 23/10/2000"
 /rpt_family="SINE/Alu"
 2849..3156
 /note="AluSx; RepeatMasker predicted
 23/10/2000"
 /rpt_family="SINE/Alu"
 3606..3663
 /note="(CCCCCG)n; RepeatMasker predicted
 23/10/2000"
 /rpt_family="Simple_repeat"
 3766..3793
 /note="GC rich; RepeatMasker predicted
 23/10/2000"
 /rpt_family="Low_complexity"
 3874..3908
 /note="GC rich; RepeatMasker predicted
 23/10/2000"
 /rpt_family="Low_complexity"
 5288..69224
 /gene="K1AAL426"
 /note="synonym: CRAMP 1 like"
 join(5288..5397,5955..6301,11954..12007,17352..17482,
 23528..23691,29140..29223,32455..32503,34063..34273,
 35284..35356,35555..35697,43780..43897,44349..44472,
 44552..44707,46535..46607,47142..48308,49423..49620,

```
51202..51379,53806..53987,54077..54151,56373..56454,
57389..57493,57738..57916,58651..58716,59278..59509,
60317..60481,61980..62125,65197..69222)
/ gene="KIAA1426"
/ note="similar to Unigene clusters HS.12849 and
HS.117900"
CDS
Join(5288..5397,5955..6301,11954..12007,17352..17482,
23528..23681,29140..29223,32455..32503,34063..34273,
35284..35356,35558..35697,43780..43897,44349..44477,
44552..44707,46535..46607,47142..48308,49423..49620,
51202..51379,53806..53987,54077..54151,56373..56454,
57389..57493,57738..57916,58651..58716,59278..59509,
60317..60481,61980..62125,65197..65361)
/ gene="KIAA1426"
/ note="Peptide sequence identity with GP.V13674.1;
Structure of CDS predicted by GENSCAN/genewise; last
coding exon was not predicted by GENSCAN but supported by
cDNA emb1:AB037847"
/ codon_start=1
/ product="unknown"
/ protein_id="AAK61288.1"
/ db_xref="GI:14336760"
/ translation="MEFVPRAGEPALPSRGGSLCPPPGALGAGWGRMTVKLGD
CGSEDDKLLKRAADEESLEGEAGADAABESSGKTRDEKTPRAGADGPAPPGA
PQAPSPQSPQDQHFLRSSVRPQSKRPKDPFSAVGNAGSGPRKDRHLPKAO
LMLSTSRVFGSSSRNLGSSGKEEGKVRWESWSTEDKNTPEFGLYEHKGD
FEATQNNIALYKKKKPASMVKNEQVRFYRTWKI.TKYIDPHVFSRGLKSSQ
BLYGLICVGEURKXIGGCMDDKNATLKNELIQVGHHTWGKSYFTFYFISSMIDGKPE
FOTLCMLEDAQGLSDERSFCQNTDVLPSGGVVGTCASIRGRTYEASVVAHLATS
LSGDWAGGKKREKRETRKLEQHGTSKERFLSAPMERSEARLLNSVFCISGVNVV
PASAPLQNSFRATVYKQNRNLRIKPMCRALKKLCDDGLSDEEDOKPVRLLPLKVP
IELOPRNHAWAQSLQNRNLRFKOFQVHYALSSHDAAVRRLESREHWA.VL
YLRDRPTCVQAVGMSMI.VELHRKYSLLEFLKQKWLHEHPDLNSQCGPSLITG
ORKTLERQLODCSAPQEKVTLHLFPGENCTLTPGVARVHSAFCTVHWQEGG
RCKSADAHVLPQAIIIGSQGGTARGQVKCPRSAGKGVGRPPPADALQSSGE
SSPSEDEGLSLDLPYLKSCODLI.VPEQCRCADTRPGSEPPGLGAASPEVLAPVS
KEAADLAPGSPRPGELLIDVCTKDLADAPAELOEKSPAGPPSPQOPAPRPK
EVPASRLAQOLREBGNLQTSLSLTAEVILMMGKPSKLOLEYDMLGFRQDPPRGLS
PTALHKLRLSLCLLKLJSTEVPNKLVPQGLPRRVGPSAGNRNVOICLSIKAGGKVL
LPGVETPEPQAPLRSRLRVKVGSHSRDTRALEANTISTASVRPAQEEQSMTPFG
KVTVSRSPQRNASLRKSTFPFSSAPCSSGLNRPRLPVLVPQSPSTGNSDSG
GLFAVPTLLPNSRSHGLFSPKAEALFTFQHLNSISMQSDFFLPKPKLRNHLRP
LVQRTLLPNSQHNVCFSILNSVTVGRGSRFPQIOSLTKAALSRPLVPKVL
PQATSHLSAIDIAATAGILSGNPLPALDTGSLGSPSSDEVTCAISQDSTGTH
QDQDTLTVGSDPFSVPSRPEQEPVADSFQGSVLSSELKPAQLONGLSLPSSS
ESSRSLSPDVSALLDISUPPEPDALSQEPATHISDSLIETAISSGQYGEVPLS
PAKLNGSDSKSLPSSSPQPHWIASPTHDPQWPSDSDTSSLSLFAFSFISPEKR
XMLPTPIGTNSGTSLLGSLDGNRDSFVRSRLADVAEVVDSOLVCMNENSDIYS
RFNDLAQELSLTAEPGRREALFDGGGGGPAVSDLSQ"
5511..6410
/ gene="KIAA1426"
/ note="CpG island; CRAMP1L"
/ evidence_not_experimental
5625..5700
/ note="GC rich; RepeatMasker predicted
23/10/2000"
/ rpt_family="Low_complexity"
5722..5828
/ note="G-rich; RepeatMasker predicted
23/10/2000"
/ rpt_family="Low_complexity"
5751..5841
/ note="(CCGGG)n; RepeatMasker predicted
23/10/2000"
/ rpt_family="simple_repeat"
6106..6171
/ note="GC rich; RepeatMasker predicted
23/10/2000"
/ rpt_family="Low_complexity"
7716..7810
/ note="MIR; RepeatMasker predicted
23/10/2000"

/rpt_family="SINE/MIR"
complement(8132..8438)
/ note="AluY; RepeatMasker predicted
23/10/2000"
/rpt_family="SINE/Alu"
9506..9625
/ note="AluSp/q; RepeatMasker predicted
23/10/2000"

Query Match 47.4%; Score 1751.4; DB 9; Length 270150;
Best Local Similarity 91.7%; Pred. No. 1.7e-247; Indels 174; Gaps 2;
Matches 1946; Conservative 0; Mismatches 1;

QY 1734 ACCCAATAGACATCCCACTCTGGGAGGCGCCCGTCCCGGGGCGATCATCCACGT 1793
DB 159584 AGCCAATAAGACATCCCACTCTGGGAGGCGCCCGTCCCGGGGCGATCATCCACGT 159643
QY 1794 GTATGGCATGACAGCAGTGCAGGGCGGCGCAGAGCTTCATCCCTACTGCTCCATGGC 1853
DB 159644 GTATGGCATGACAGCAGTGCAGGGCGGCGCAGAGCTTCATCCCTACTGCTCCATGGC 159703
QY 1854 CCAGCCCAAGCTATGCTTCCATGGCACCCGCGATGCCGTGAAGTTCTTTTGTCTCGTG-- 1911
DB 159704 CCAGCCCAAGCTATGCTTCCATGGCACCCGCGATGCCGTGAAGTTCTTTTGTCTCGTG-- 159763
QY 1912 ----- 1911
DB 159764 AGGTGAGGCTGGGCGCCCTCTCTGCCATCCACATCCCATGCCAGTGGCGCGCCCTCC 159823
QY 1912 -----CCAGGGAACGTGCTGGCCACCCCTG 1935
DB 159824 CCCAGGAGCGCGTGTCTGTAATCGTTCTGCCATCCACAGGGAACGTGCTGGCCACCCCTG 159883
QY 1936 AATGGCAGTGTGCTGGACAGCCAGCCAGCGAGGCCCTGGGCGCAGTGCCTCGCTCGGAG 1995
DB 159884 AATGGCAGTGTGCTGGACAGCCAGCGAGGCCCTGGGCGCAGTGCCTCGCTCGGAG 159943
QY 1996 GTCAGGGCCAGAGCTGCGGAAACGTGCTGTGTGTGAGCGCGGGGAGGGGTACATCGAC 2055
DB 159944 GTCAGGGCCAGAGCTGCGGAAACGTGCTGTGTGTGAGCGCGGGGAGGGGTACATCGAC 160003
QY 2056 TTCCGCAATT----- 2064
DB 160004 TTCCGCAATTGGTGTGAGCGGGGCCAGGACAGGGGTGTGAGTTTGGCGGGGAGGACCTTG 160063
QY 2065 -----CGAGACGAGAGACAGACAGCAGACGAGAGGAGGGCGC 2099
DB 160064 CCCTCACTCTGCTTTGCGCGAGAGACGAGAGAGACGAGAGAGCGGAGGGGGCGC 160123
QY 2100 AGGGACATGAGCCAGTGAAGCCCGTGTGTCCAAAGCGAGAGCGCAGTCAATCATCGT 2159
DB 160124 AGGGACATGAGCCAGTGAAGCCCGTGTGTCCAAAGCGAGAGCGCAGTCAATCATCGT 160183
QY 2160 GTGCGAGTGTCTTACACCCCGAGTGAAGCTGCTGCTGCTGGCCCGACCTGTACAT 2219
DB 160184 GTGCGAGTGTCTTACACCCCGAGTGAAGCTGCTGCTGCTGGCCCGACCTGTACAT 160243
QY 2220 AGGACCCCGACCACTGACCCCGCGCGCGGGGTAGCCAGCAGCGCGCGCGCGC 2279
DB 160244 AGGACCCCGACCACTGACCCCGCGCGCGGGGTAGCCAGCAGCGCGCGCGCGC 160303
QY 2280 CCCTCTTAACTCTCAACCTGACCTTTCACCTGAGTGTGCGCCCTCCAGCGGGCAGG 2339
DB 160304 CCCTCTTAACTCTCAACCTGACCTTTCACCTGAGTGTGCGCCCTCCAGCGGGCAGG 160363
QY 2340 GAGTGGGGGATCGGATCAGCTGGGAGGAGGAGGGGTGCTTCCACCGAGGGGA 2399
DB 160364 GAGTGGGGGATCGGATCAGCTGGGAGGAGGAGGGGTGCTTCCACCGAGGGGA 160423
QY 2400 AGATGCTCTCGGACAGTGTTCGCCGCGAGCTCTCTGCGCAGCTTCCAGCCAGAGTCTCTCA 2459
DB 160424 AGATGCTCTCGGACAGTGTTCGCCGCGAGCTCTCTGCGCAGCTTCCAGCCAGAGTCTCTCA 160483
```


Db	5328	CCAGGCCCAGCTATGCTTCCATGGGCACCGCATGCCGTGAAGTTCTTTGTCTCGGTGCC	5387
QY	1912	-----	1911
Db	5388	AGGTGAGGCTGGGCCCTTCCTGCCATCCACATCCCTTCATGCGCATGGCGCGCGCCTCC	5447
QY	1912	-----CCAGGAAAGTGTCTGCCCACCCCTG	1935
Db	5448	CCAGGAGCCGCTGTCTCTGAATCGCTTCTGCCATCCACAGGAAAGTGTGCGCCACCCCTG	5507
QY	1936	AATGCAAGTGTCTCGACAGACCCACGACCGAGAGGCCCTTGGCCAGCTGCCCTTGCCTCGAG	1995
Db	5508	AATGGAGTGTCTGGACAGCCACCGAGAGGCCCTTGGCCAGCTGCCCTTGCCTCGAG	5567
QY	1996	GTGAGGGCCAGAAAGCTTCGGAAAGTGTGTGTGTGAGCGCGCGGAGGGGTACATCGAC	2055
Db	5568	GTGAGGCCAAGAAGCTCGGAAAGTGTGTGTGTGTGAGCGCGCGGAGGGGTACATCGAC	5627
QY	2056	TTCCGCATT-----	2064
Db	5628	TTCCGCATTGTGAGCGGGGCCCAGGGACAGGGCTGAGTTGGCGCGGGGGAGCCCTGG	5687
QY	2065	-----GGNAGCGAGAGAGCGACGAGAGGAGGCGC	2099
Db	5688	CCGTCACTCTGCTGTTTGGCCGAGAGAGCGAGAGGACGACGAGCGGAGGAGGCGC	5747
QY	2100	AGGGACATGAGCCAGAGTGAAGCCCGCTGCTGTCCAAAGCGAGCGGAGTCAACATCATCGT	2159
Db	5748	AGGGACATGAGCCAGAGTGAAGCCCGCTGCTGTCCAAAGCGAGCGGAGTCAACATCATCGT	5807
QY	2160	GTGCAAGTGTCTTACACCCCGAGTGAAGTGTGTGCCCTTGCCTTGGGCCACCTGTACAT	2219
Db	5808	GTGCAAGTGTCTTACACCCCGAGTGAAGTGTGTGCCCTTGCCTTGGGCCACCTGTACAT	5867
QY	2220	AGGACCCCGACACCTGACCCCGCCCGCCCGCGGGGTAGCCAGCGAGCGCGCGCGC	2279
Db	5868	AGGACCCCGACACCTGACCCCGCCCGCCCGCGGGGTAGCCAGCGAGCGCGCGCGC	5927
QY	2280	CCCTCTTTAACTCTCAAGCTGCAGCTTTTCACTGAGTCTGCGCCCTCCAGCGGCGCAGG	2339
Db	5928	CCCTCTTTAACTCTCAAGCTGCAGCTTTTCACTGAGTCTGCGCCCTCCAGCGGCGCAGG	5987
QY	2340	GAGTGGGGATGCGGATTCAGCTGGGAGGAGGGAGGGGTCTTCCACCCGAGGGGA	2399
Db	5988	GAGTGGGGATGCGGATTCAGCTGGGAGGAGGGAGGGGTCTTCCACCCGAGGGGA	6047
QY	2400	AGATGCTCTCGGACAGTTTCCCGGCAGCTCTGCGCAGCTTCCAGCCGAGAGTCTCA	2459
Db	6048	AGATGCTCTCGGACAGTTTCCCGGCAGCTCTGCGCAGCTTCCAGCCGAGAGTCTCA	6107
QY	2460	AGTCCAGGACCTTGGGCCACGCGAGGAGAGATCCGAGTGGTCTGCTCTACCCCTG	2519
Db	6108	AGTCCAGGACCTTGGGCCACGCGAGGAGAGATCCGAGTGGTCTGCTCTACCCCTG	6167
QY	2520	GGCTCTCTATCCCGACGACCCCTGAGGAGGAGGGGCTCCCGCCGCGAGGCTGCT	2579
Db	6168	GGCTCTCTATCCCGACGACCCCTGAGGAGGAGGGGCTCCCGCCGCGAGGCTGCT	6227
QY	2580	GCCTTGGGCCACCTCTGATGCTGTCTATGGGGCCACCGCTGCTCTGGGCGCTCACTC	2639
Db	6228	GCCTTGGGCCACCTCTGATGCTGTCTATGGGGCCACCGCTGCTCTGGGCGCTCACTC	6287
QY	2640	TGCTAGGGAGCTGGGCCACGACCTAGGCTTTGGCCAGGAGGTGGGCTCAGGCTGCC	2699
Db	6288	TGCTAGGGAGCTGGGCCAGGACCTAGCTTTGGCCAGGAGGTGGGCTCAGGCTGCC	6347
QY	2700	CAGGTGCTGCACCCGAGCGCGCTTCTCTGGGGCCCTCCCGTGTGAAGCCCTCTATCCT	2759
Db	6348	CAGGTGCTGCACCCGAGCGCGCTTCTCTGGGGCCCTCCCGTGTGAAGCCCTCTATCCT	6407
QY	2760	GTCTGTCCCCACCCAGCTGTCCCTTGCACGAGGAGCTGGCATAAAGCACGAGGCCCGG	2819
Db	6408	GTCTGTCCCCACCCAGCTGTCCCTTGCACGAGGAGCTGGCATAAAGCACGAGGCCCGG	6467

RESULT 13
AF262046
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Db 4295 GCAGGAACAGATGCTTTTCCCATCAGGGAGAGAGCGCTTGAGACAATCTCTCCAGCA 4395

Qy 2428 GCTCCTGCCAGCTTCAGCCAGATCC 2456
||||| ||||| ||||| |||||
Db 4355 GCCCTGTCTGTGACCACCAATGAAGTCC 4383

RESULT 14

AB005662 4173 bp mRNA linear ROD 29-OCT-1999

LOCUS Mus musculus mRNA for JNK/SAPK-associated protein-1, complete cds.

DEFINITION AB005662.1 GI:6141548

ACCESSION JNK/SAPK-associated protein-1.

VERSION Mus musculus (house mouse)

KEYWORDS Mus musculus

SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Ito, M., Yoshioka, K., Akechi, M., Yamashita, S., Takamatsu, N.,
Sugiyama, K., Hibi, M., Nakabeppu, Y., Shiba, T. and Yamamoto, K.I.,
JSAPI, a novel jun N-terminal protein kinase (JNK)-binding protein
that functions as a Scaffold factor in the JNK signaling pathway
Mol. Cell. Biol. 19 (11), 7539-7548 (1999)

MEDLINE 99455010

PUBMED 10523642

REFERENCE 2 (bases 1 to 4173)
Yoshioka, K.
Direct Submission
Submitted (09-JUL-1997) Kateuji Yoshioka, Cancer Research
Institute, Kanazawa University, Molecular Pathology; 13-1
Takaramachi, Kanazawa, Ishikawa 920-0934, Japan
(E-mail:kateuji@kenroku.ipc.kanazawa-u.ac.jp, Tel:81-76-265-2757,
Fax:81-76-234-4517)

TITLE Location/Qualifiers
1. .4173
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="JSAF1"
/tissue type="brain"
107. .4024
/codon_start=1
/product="JNK/SAPK-associated protein-1"
/protein_id="BAA85974.1"
/db_xref="GI:6141549"

CDS
/translation="MMETITDGGGVVYQDDYCSGVSMSRVSLAGSIYREPERLRI
HCYEVEVKMLPLVNLVDLSNLESGDEVELELRDNELLOQLTYERESEKALRK
QAEEKIFPEDALEQEKKELQIQVEHYEFOTRLELKAKNVADQISRLRESEMKKE
YNALHQRHTMIQTVEHERSKMOOVGGSGQTSSIPGRKERPTSLNVPFLADGM
PNDEMSEQSQAATPTGTGTSNTPTSSPSVSPAAPELNESLQPLGDIVSYTKNNKQ
AREKNSRNWEQVTQMERNVISGSSDEWSVDODIIDSTPELDVCPTERLETGS
PTQGNKAFTNDLSLYHELTAGSEVIGDVDEGALLGEFAKQAVLENRIKLEIE
LQENGLKETNALNVNDLIATKDQLSGEVLKGLEAAKQAVLENRIKLEIE
ELRKLQRAVRWTEIMIRASREHPSVOEKKSITMQCFSELSSSSPPAKRSYPVN
IHVKPPTAAGFSQRSHALCQISASRFLEFPDDCTSSARREQRRQYQVRHVHR
NDDRLQGWSLPARYKQLSPNGQEDTRMNKPVPYCRPLVSKKETPEADATSS
NLSUGMPHEESSNGEPKPVDRDLTCDREGEGEPKTHSPSEKKKAKETPEADATSS
RWYLITITTTSKVVIDANPGTIDVOTVCNAHLCTGISPAASDSDPPEGMFLD
SDNPEDSGADGLAGITLVGCATCNVPSNRSSRGDTPLVDKGQGVATTANGKN
NGSTIVEATEVEVPDGPSEEAATVPDGLTHEVTDTPATPSSSTQPASENGSES
PQTIVQPVQPSLEGTLTSTSSAAPTMWLGAQNGLYHRSVAHNKKCUHSILKLDVSL
SUVHVGKRVVALADGLTIATHRGEDGMDLSNHLMDLGHPHHSIRCMAYVNDKRW
GYKNXKVIQPMQIOIKSFDAHPRESQVRQLAWIGDGVVWSIRLDSLTRLYHAHTH
QHLDVIDEPYVKMTGLTKGFSPVRIATLLIAGNRLWGVTGWGWTISIPLTETVL
RDQGLLGRLANKRTSPTSCGERTPGCI IHVYGDDSDKAASSFIYCSCMAQQLCFEHG
RVAKFFVYSFGNVLATINGSVLSPSGPGPAADAECQKLNALVLUSGGGYID
FRIGDGEDDETCCAGDVNQTKPSLSUKASRHSIIWQVSYTFE"

ORIGIN

Query Match 41.5%; Score 1531.6; DB 10; Length 4173;
Best Local Similarity 82.3%; Pred No 9, le-215;

Matches 1785; Conservative 0; Mismatches 379; Indels 6; Gaps 2;			
Qy	148	GATGGGTAGGAGCCAGGGTGTGTCCTCCACGGCGCTCCCTGCTCCCTGCGAGTCAGTCCC	207
Db	1985	GACGGGAGGTGTCAGGCTGTGGGTGGAGCTGCTGCTGCAAGTACAAGCAGCTGAGCCCC	2044
Qy	208	AACGGGGCCAGGAGGACACCGGATGAAGAACGTGCCGGTCCGGTGTACTGTCCGGCCCT	267
Db	2045	AATGGAGGCCAGGAGACACCCGATGAAAAATGTGCTGCTGCTGTGTACTGTGCGCCCT	2104
Qy	268	CTGTGTGAGAGGACCCACCATGAAGCTGTGGTGTGCGCGGGCGTCAACCTCAGCGGG	327
Db	2105	CTGTGTGAGAGGACCCCTTCACAAAGCTGTGGTGTGCTGTGCTGTGTGTCAACTGTGGG	2164
Qy	328	TGGAGGCCCAATGAGGACACGCTGTGGGAATGGAAGTCAAGCCAGCGCCAGGCCGGATCCC	387
Db	2165	TGGAAGCCACATGAAGAGACTCTAGCAATGGAACCCAAAGCCTGTACAGGTGAGACCCCT	2224
Qy	388	CTGACCTGCGACCGGAAGGAGACGGCGAGCCCAAGAGCGCCACACGTCTCCCGAGAAG	447
Db	2225	CTGACCTGTGACCCGGGAAGGAGAGCCAAACCCAGAGCAACACCCCAATCACTGTAGAG	2284
Qy	448	AAGAAGGCCAAGGAGCTCCCTGAAATGACGCACTCCAGCCGGGTGTGGATCCTGACC	507
Db	2285	AAGHAGGCCAAAGGAACCCCTTGAGCGAGATGCTACTCCAGTCCGGTATGATCCTCACC	2344
Qy	508	AGCACCTTGACACACAGCAAGGTGTGATCATPAGACGCAACACGACCGGGCAAGTGTG	567
Db	2345	AGCACCTTGACAAACAGCAAGGTGTGATCATTTGATGCGCAACACGACCGGCAAAATTGT	2404
Qy	568	GACAGTTTACCGTGTGCAACCGCACGTCGTGTGATCTCCAGATCCCGCGGCCAGC	627
Db	2405	GATCAGTTTACAGTGTGAATGCCACGTCCTGTGTATCTCCAGATTTCTTCGCGCCAGT	2464
Qy	628	GACACGACTACCTTCCCGGGAGATGTTCTTGGACAGCGAGTGAACCCAGAGACCCG	687
Db	2465	GACAGTGAATATCCCTCTGGGAGATGTTCTTAGACAGTATGTGAACCTTGAAGATTCA	2524
Qy	688	GGCGAGATGGCGTGGCGCGGTATCACCTGTGTGGGTGTGTCACCCGCTGCAACGTG	747
Db	2525	GGTGTGATGGTGTGCTGGCTGGCATCACCTGTGTGGGTGTGCTTACCCGCTGCAATGT	2584
Qy	748	CCGCGGAGCAATGTCTCTCCGAGGGGACACCCAGTGTCTAGACAAGGGGACGGGAG	807
Db	2585	CCACGTAGCAATGTCTTCTACAGAGAGACACCCAGTACTGGAACAAGGGGACGGGAT	2644
Qy	808	GTGGCCACCATCGCCAAACGGGAAGGTCAACCCGTCCCAGTCCACAGAGAGGCCACAG	867
Db	2645	GTGGCGACCACTGCCAATGGGAAGGTCAACCCGTCCCAGTCCACAGAGAGGCCACAG	2704
Qy	868	GCCAGGAGGTGCAGACCTTGGGCCACGAGCGCAGAGACAGCCACATTTGGCGCCGGG	927
Db	2705	GCCACAGAGGTGCCAGCCCTGTCTCCAGCGAGTCAGAGCAACGACAGTCCGCGCCGGG	2764
Qy	928	CCTCTCACAGACAGCTTCTCACTGACCCAGCCCGGACCCCTCTCTTGGCCCGCCAGCCT	987
Db	2765	CCTCTCACAGAGATGTCTTTACTGACCCAGCACCCACCCCACTCTCCAGACCCAGCCT	2824
Qy	988	GGCAGCGAAGCCGGGCCAGAGCCTTGACAGCAGCAGCACACGGCCAGAGCCAGGCCAGC	1047
Db	2825	GCCAGTGAATGGGTCAGAATCCAATGGCACCAATGTGTACAGCCTCAGGTGGAGCCAGT	2884
Qy	1048	GGGGACCCACGGGAGCAGGACGTGTGTCACCCACCATGTGGTGGAGGCCAGAAC	1107
Db	2885	GGGGAACTCTCAACAAACACAGTAGCGTGTGACCCATGTGTGGTGTAGAGCCCAAGAT	2944
Qy	1108	GGCTGGCTCTATGTGCACTCGGCTGTGGCCAACTGGAAGAAGTGCCTGCACTCCATCAAG	1167
Db	2945	GGCTGGCTCTATGTGCAATTCAGCGGTAGCCAACTGGAAGAAGTGTCTGCACTCCATCAAG	3004
Qy	1168	CTGAAGGATTTGTGTCAGCCTGGTGCATGTCAAAAGGCGGTGTGTGTGGTCTTGGCG	1227
Db	3005	CTAAAAGACTCTGTGCTGAGCCTGGTGCATGTCAAAAGGCGGAGTGTGTGTCTTGCA	3064

Qy	1228	GACGGACCTGGCCATCTTCCACCGTGGTGAAGATGGCCAGTGGGATCTGAGCAACTAT	1287
Db	3065	GATGGGACCTTGGCTATCTTCCATCTGTTGGAGAGATGGCCAGTGGGACCTGAGCAACTAC	3124
Qy	1288	CACCTAATGACCTGGGCCACCCGACACTCCATCCGTGCAATGCGTGTGTGTGTACGAC	1347
Db	3125	CACCTAATGACCTGGGCCACCCACACTCCATCCGTGCAATGCGTGTGTGTGTGTGAATGAC	3184
Qy	1348	CGCTGT	1407
Db	3185	CGAGTTTGGT	3244
Qy	1408	GAGAAAGTCAATTTGACGCCACCCGCGCGGAGAGCCAGGTGCGGCACTGCGCGTGGATC	1467
Db	3245	GAGAAATCAATTTGATGCCACCCAAAGCGGGAAGACAGGTACGTACGTGCGCTGGATC	3304
Qy	1468	GGGATGCGGTATGGGTGTCCATCGGCTGGAATCCACCTGAGGCTCTACATGCAAC	1527
Db	3305	GGTATGAGGTGTGGGTCTCTATTTCCGTTGGATTTCTACCTTTCGCTCTACATGCTCAC	3364
Qy	1528	ACGCACAGCATCTACAGGACGTGGACATTTGAGCCCTACGTACGACAGATGCTAGGCACT	1587
Db	3365	ACCCACAGCACCTGCAAGGATGTGGACATTTAGCCCTATGTTAGCAAGATGCTAGGAACC	3424
Qy	1588	GGCAAGCTGGGTCTTCTCTCTGATCGCATCACGGCCCTGTGTCGCGGGGAGCCGGCTC	1647
Db	3425	GGCAAGCTGGGTCTTCTCTCTGATCGCATCACGGCTTACTCATTTGACGGAACCGCTCG	3484
Qy	1648	TGGGTGGGCAACGGGAGTGTGTATCTCCATCTCCCTGACAGACATGTTGTGTGTGTGT	1707
Db	3485	TGGGTGGGCACTGGCAATGGGGTGTGTATCTCCATCTCCCTTGTGACTGAGACTGTGTGT	3544
Qy	1708	CACGAGCCAGCTCTTGGGCTCGAGCCCAATAGACATCCCCACCTCTGGGGAGGGC	1767
Db	3545	CATGAGCCAGCTCTTAGGGCTCGAGCCCAACAGACATCCCCAACATCTGGGGAGGGG	3604
Qy	1768	CCCCGTCCGGGGGATCATCCAGTGTATGGCCATGACAGCAGTGAACAGGGCGGCCAGC	1827
Db	3605	ACCCGCCAGGGGGATCATCCATGTGTATGGGACGACAGCAGTGAACAGGGCGGCCAGT	3664
Qy	1828	AGCTTATCCCTTACTGTCTCATGCGCCAGGCGCCAGTATGTTTCCATGGGGACCGCGAT	1887
Db	3665	AGTTTATCCCTTACTGTCTCATGCGCACAGGCTCAGCTTGTCTTCCATGGGACCGCGAT	3724
Qy	1888	GCCGTGAAGTCTTGTCTCGTGCAGGGAACGTGTCGGCCACCTGAAATGCGCAGTGTG	1947
Db	3725	GCTGTCAAATCTTGTCTGTGCGCAAGTGTGTGGCCACTCTCAATGGCAGTGTG	3784
Qy	1948	CTGGACAGCCACGCGGCGCTGGGCGAGCTGCGCCCTGCTCGAGGTTCGAGGGCCAG	2007
Db	3785	CTAGACAGCCCATCAGAGGGCCCTGGGCTGTGTCACCCGCTGCAGATGCTCAGGGCCAG	3844
Qy	2008	AAGTTCGGAACGT	2067
Db	3845	AAGTTGAAGATGCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3904
Qy	2068	GACGAGGAGGACGACGAGACGAGAGGAGGCGCAGAGGACATCAGCCAGGTGAAGCCCGT	2127
Db	3905	GACGAGGAGGATGATGAAGT	3964
Qy	2128	CTGTCTCAAGGCGAGCGAGTCAATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2187
Db	3965	TTGTCTCAAGGCTGAGCGCAGCCACATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4023
Qy	2188	AGCTGTGCCCTTGT	2247
Db	4024	AGACCCCTGTCTTGT	4083
Qy	2248	GGCCGCGGGGTAGCCAGCCAGCGCGCGCGCGCGCGCTTCTTAACTCTCTCAACTGAGCT	2307
Db	4084	GTTCCTTGGGGAGCCA-----GGTTCGTCCATCCCTTTTAACTCTCAACTGAGCT	4138

[illegible]

ORIGIN

Db 2926 CTAAGAGACTCTGCTGAGCCTGGTCACTGTAAGAGCCGAGTCTGAGTCTTTGCA 2985
QY 1228 GACGGGACCCCTGGCCATCTTTCCACCGTGGTGAAGATGCCAGTGGGATCTGAGCAACTAT 1287
Db 2986 GATGGGACCCCTGGCTATCTTCCATCGTGGAGAGATGGCCAGTGGGACCTGAGCAACTAC 3045
QY 1288 CACCTAATGACCTGGGGCCACCCGACCACTCCATCGCTGCATGGCTGTTGTGTACGAC 1347
Db 3046 CACCTAATGACCTGGGGCCACCCGACCACTCCATCGCTGCATGGCTGTTGTGAATGAC 3105
QY 1348 CGCGTGTGTGTGCTCAAGAACAGGTGACGTCACTCCAGCCCAAGACCACCATGCAGATA 1407
Db 3106 CGAGTTTGGTGTGCTACAAGAACAGGTGATGTATCCAGCCCAAGACAAATGCAGATT 3165
QY 1408 GAGAAAGTCATTGTAGCCCCACCCCGCGGCGAGAGCCAGGTGGGCGAGCTGGCGTGGATC 1467
Db 3166 GAGAAATCATTTGATGCCACCCCAAGCGGGAAGCCAGGTACGTCACTGGCTGGCTGGATC 3225
QY 1468 GGCATGCGCTATGGGTGCTCATCGGCTGACTCCACCTGAGGCTTACCATGCACAC 1527
Db 3226 GGTGATGGAGTGTGGGCTTCTATTGCTTGGATTCTACCCCTGGGCTTACCATGCTCAC 3285
QY 1528 ACGCACAGCATCTACAGAGCTGGACATTTAGCCCTACGTACGCAAGATGCTAGGCACT 1587
Db 3286 ACCCACAGCACTGCGAGGATGGACATTTAGCCCTATGTTAGCAAGATGCTAGGAACC 3345
QY 1588 GGCAGCTGGGTTTCTCCTGTAGCATACGGCCCTGCTTGTGCGGGGAGCGGCTC 1647
Db 3346 GGCAGCTGGGCTTCTCCTGCGGCATCACAGGCTTACTCATTTGCAGGCAACCGTCTG 3405
QY 1648 TGGTGGGACCGGCAACGGAGTGTTCATCTCCATCCGCTGACAGACTGTGTCCTG 1707
Db 3406 TGGTGGGCACTGGCAATGGGTTTGTCTATCTCCATCCGCTTACCTGAGACTGTGTCCTG 3465
QY 1708 CACGAGGCCAGCTCCTGGGCTCCGAGCCAAAGACATCCCCACCTCTGGGAGGGC 1767
Db 3466 CATGAGGCCAGCTCCTAGGGCTCCGAGCCAAAGACATCCCAACATCTGGGAGGGG 3525
QY 1768 GCGGTCCTGGGGGATCATCACTGTATGGCGATGACAGAGTGACAGGGCGGCCAGC 1827
Db 3526 ACCGCCCCAGGGGATCATCATCTGTATGGGAGCAGCAGGTGACAAAGGCCGCCAGT 3585
QY 1828 AGCTTCATCCCTACTGCTCCATGCGCCAGGCCAGCTATGCTCCATGGGCAACCGCAT 1887
Db 3586 AGTTTCATCCCTACTGCTCCATGCGCAGAGGCTCAGCTTGTCTCCATGGGCAACCGTAT 3645
QY 1888 GCGGTGAAGTCTTTGTCTCGGTGCGAGGGAACGTGTGGCCACCTGAAATGGCAGTGTG 1947
Db 3646 GCTGTCAAATCTTTGTCTCTGTGCCAGGAATGTGTGGCCACTCTCAATGGCAGTGTG 3705
QY 1948 CTGGAAGCCAGCCGAGGGCCCTGGGCGAGCTGCCCCCTGCGCTCGAGGTGAGGGCCAG 2007
Db 3706 CTAGACAGCCCATCAGAGGGCCCTGGGCTGTGTGACCCGCTGCAGATGCTGAGGGCCAG 3765
QY 2008 AAGCTGCCGAGCTGCTGTGTGAGCGCGGGAGGGCTACATCGACTTCCGCAATGGA 2067
Db 3766 AAGTTGAAGATGCACTGGTGTGAGTGGTGTGAGGTTTCAATTGACTTCCGTATCGGA 3825
QY 2068 GACGAGAGGACGACGAGACGAGAGGGCGGAGGGGACATGAGCCAGGTGAAGCCCGTG 2127
Db 3826 GACGAGAGGATGATGAACACTGAGGNACTGTGCCGGGACGTGAACAGACAAAGCCCTCG 3885
QY 2128 CTGTCCAGGAGAGCGCAGTCAATCATCTGTGTGGCAGGTGTCTACACCCCGAGTGA 2187
Db 3886 TTGTCCAGGCTGAGCGCAGCCACATCATCTGTGTGGCAGGTGTCTTACACCCCTGAGTGA 3945

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 11:03:47 ; Search time 59 seconds
(without alignments)
3117.600 Million cell updates/sec

Title: US-10-019-495-9
Perfect score: 3473
Sequence: 1 MKNVFPVYCRPLVEKDPTM.....VLSKAERSHIIWQVSYTPE 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3473	100.0	651	3 AAB29652	Aab29652 Human mem
2	3468	99.9	1066	4 AAM41910	Aam41910 Human pol
3	3468	99.9	1069	3 AAB42717	Aab42717 Human ORF
4	3465	99.8	1014	4 AAM40124	Aam40124 Human pol
5	3369	97.0	1069	4 ABG09904	Abg09904 Novel hum
6	3191	91.9	1305	3 AAB12875	Aab12875 Murine JN
7	3191	91.9	1314	3 AAB12876	Aab12876 Murine JN
8	3191	91.9	1336	3 AAB12878	Aab12878 Murine JN
9	3191	91.9	1337	3 AAB12877	Aab12877 Murine JN
10	2757.5	79.4	809	4 ABG09902	Abg09902 Novel hum
11	1821.5	52.4	1307	6 ADA45192	Ada45192 Murine JN
12	1814.5	52.2	1311	6 ADA45190	Ada45190 Human JLP
13	1732.5	49.9	1277	4 ABB11784	Abb11784 Human sec
14	1467	42.2	1165	6 ADA45195	Ada45195 Murine JL
15	1446.5	41.6	828	2 AAW64473	Aaw64473 Human sec
16	1446.5	41.6	828	4 AAB90749	Aab90749 Human ECL
17	1152	33.2	1195	4 ABB63694	Abb63694 Drosophil
18	943	27.2	734	4 AAB95634	Aab95634 Human pro
19	807.5	23.3	235	3 AAB54174	Aab54174 Human pan
20	675.5	19.5	1000	6 ADA45196	Ada45196 Murine JL
21	623.5	18.0	200	4 AAB94656	Aab94656 Human pro
22	511	14.7	118	5 ABB89836	Abb89836 Human pol
23	462	13.3	260	4 ABG09899	Abg09899 Novel hum
24	462	13.3	592	4 ABG15607	Abg15607 Novel hum
25	460.5	13.3	766	7 ADB79917	Adb79917 Human put

ALIGNMENTS

RESULT 1
AAB29652
ID AAB29652 standard; protein; 651 AA.
XX AC AAB29652;
XX XX
DT 23-FEB-2001 (first entry)
XX XX
DE Human membrane-associated protein HUMAP-9.
KW Human membrane-associated protein; HUMAP; transgenic organism;
KW drug screening; cell signalling modulator; agonist; antagonist;
KW cell differentiation modulator; cell proliferation modulator;
KW cell proliferative disorder; cancer; cell differentiation disorder;
KW developmental disorder; cell signalling disorders; endocrine disorder;
KW hyperpituitarism; hypothyroidism; hyperparathyroidism; infection;
KW pancreatic disorder; diabetes mellitus; immunological disorder;
KW hereditary neuropathy; gonadal steroid hormone associated disorder;
KW infertility.
XX XX
OS Homo sapiens.
XX XX
PN WO200065054-A2.
XX XX
PD 02-NOV-2000.
XX XX
PF 20-APR-2000; 2000WO-US010884.
XX XX
PR 23-APR-1999; 99US-0130694P.
PR 23-JUN-1999; 99US-0140580P.
XX XX
PA (INCY-) INCYTE GENOMICS INC.
XX XX
PI Hillman JL, Bandman O, Tang YT, Lal P, Yue H, Reddy R;
PI Azimzai Y, Baughn MR;
XX XX
DR WPI; 2000-687346/67.
DR N-PSDB; AAC64282.
XX XX
PT Human membrane-associated protein, useful for diagnosis and treatment of
PT cell signaling, cell differentiation and cell proliferation disorders
PT such as cancer, and for identifying agonists and antagonists.
XX XX
PS Claim 1; Page 78-79; 99pp; English.
XX XX
CC The invention relates to 17 human membrane-associated proteins, HUMAP-1
CC to HUMAP-17 (AAB29652) and the cDNAs encoding them (AAC64274-
CC C64290). The invention also relates to expression constructs, host cells

26 442.5 12.7 1935 6 ABJ19396 Abj19396 Human int
27 442.5 12.7 2063 5 ABB90736 Abb90736 Human tum
28 442.5 12.7 2063 6 ABU54443 Abu54443 Human tum
29 436.5 12.6 1101 4 ABB63244 Abb63244 Drosophil
30 333.5 9.6 1195 6 ABU11484 Abul11484 Human MDD
31 333.5 9.6 1233 5 ABP55147 Abp55147 Neurodeve
32 333.5 9.6 1240 6 ABU54572 Abu54572 Human NOV
33 306 8.8 988 4 ABG08836 Abg08836 Novel hum
34 305.5 8.8 988 7 ADE09396 Ade09396 Novel pro
35 305.5 8.8 1129 4 ABG08838 Abg08838 Novel hum
36 265 7.6 1783 4 ABB63930 Abb63930 Drosophil
37 233 6.7 596 4 AAB92909 Aab92909 Human pro
38 191.5 5.5 91 4 AAU32873 Aau32873 Novel hum
39 186.5 5.4 119 5 ABB89345 Abb89345 Human pol
40 185.5 5.3 52 4 ABB40157 Abb40157 Peptide #
41 185.5 5.3 52 4 AAM33802 Aam33802 Peptide #
42 185.5 5.3 52 4 AAM73607 Aam73607 Human bon
43 185.5 5.3 52 4 AAM60922 Aam60922 Human bra
44 185.5 5.3 52 4 ABG55340 Abg55340 Human liv
45 185.5 5.3 52 5 ABG43477 Abg43477 Human pep

and transgenic organisms comprising a HUMP nucleic acid sequence; the recombinant preparation of a HUMP; methods of screening compounds for their ability to modulate HUMP activity or expression; and pharmaceutical compositions comprising a HUMP protein, a HUMP agonist or HUMP antagonist. The HUMP acts as modulators of cell signalling, differentiation and proliferation. A HUMP is useful for screening a compound for effectiveness as an agonist or antagonist of HUMP activity. The protein, or the identified agonist or antagonist is useful for treating a disease or condition associated with decreased or increased expression of functional HUMP. A HUMP nucleic acid is useful for screening a compound for its ability to alter expression of that particular HUMP gene. A wide variety of disease may be treated using compositions of the invention. These diseases include cell proliferative disorders (e.g., actinic keratosis, arteriosclerosis); cancer (e.g., breast, bladder, bone marrow, brain and uterine cancer); cell differentiation disorders, in particular developmental disorders (e.g., renal tubular acidosis, anaemia, Cushing's syndrome, achondroplasia, epilepsy, and muscular dystrophy); cell signalling disorders, in particular endocrine disorders such as hypothalamus and pituitary disorders resulting from lesions such as thrombosis; disorders associated with hyperpituitarism (e.g., acromegaly); disorders associated with hypothyroidism (e.g., goitre); hyperparathyroidism; pancreatic disorders such as type I or type II diabetes mellitus; infections; immunological disorders; hereditary neuropathies (e.g., neurofibromatosis); and disorders associated with gonadal steroid hormones (e.g., infertility, endometriosis, polycystic ovary syndrome, osteoporosis, Leydig cell deficiency and gynecomastia). Antibodies which specifically bind HUMP may be used for the diagnosis of disorders associated with the expression of HUMP, or in assays to monitor patients being treated with HUMP or agonists, antagonists or inhibitors of HUMP. The present sequence represents a HUMP of the invention

Sequence 651 AA;

Query Match 100.0%; Score 3473; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.1e-270;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKNVPVYCRPLVEKPTMKLWCAAGVNLGWRNEDDAGNGVKPAPGRDPLTCDEGD 60
Db 1 MKNVPVYCRPLVEKPTMKLWCAAGVNLGWRNEDDAGNGVKPAPGRDPLTCDEGD 60

QY 61 GEPKSAHTSPEKKAKELPEMDATSRVWLTLTSTTTSKVLIIDANPGTVVDQFTVCNA 120
Db 61 GEPKSAHTSPEKKAKELPEMDATSRVWLTLTSTTTSKVLIIDANPGTVVDQFTVCNA 120

QY 121 HVLICISSIPAAASDSDYPPGEMFLDSVNPDPGADGLVAGITLVGCATRCNVPNSCSR 180
Db 121 HVLICISSIPAAASDSDYPPGEMFLDSVNPDPGADGLVAGITLVGCATRCNVPNSCSR 180

QY 181 GTPVLDKGGQGEVATANGKNPSSOSTEATEATEVPDPGPEPETATILRPGPLEHVT 240
Db 181 GTPVLDKGGQGEVATANGKNPSSOSTEATEATEVPDPGPEPETATILRPGPLEHVT 240

QY 241 DPAPFPSSGPQSGENGPEPSSSTRPEPEPSSDPTGAGSSAAPTMLGAQNGWLYVHSA 300
Db 241 DPAPFPSSGPQSGENGPEPSSSTRPEPEPSSDPTGAGSSAAPTMLGAQNGWLYVHSA 300

QY 301 VANWKKLHSLIKDLSVLHVKGRLVALADGTILAI FHRGEDGQWDL SNYHLMDLGHP 360
Db 301 VANWKKLHSLIKDLSVLHVKGRLVALADGTILAI FHRGEDGQWDL SNYHLMDLGHP 360

QY 361 HHSIRCMAYVDRVWCGYKKNVHVITQPTMQIEKSFDAHPRESQVRLAWIGDGVWVSI 420
Db 361 HHSIRCMAYVDRVWCGYKKNVHVITQPTMQIEKSFDAHPRESQVRLAWIGDGVWVSI 420

QY 421 RLDSTLRLVHAHQHQLQDVITPEVYSKMLGTGKLGFSFVRITALLVAGSRLVWGTNGV 480
Db 421 RLDSTLRLVHAHQHQLQDVITPEVYSKMLGTGKLGFSFVRITALLVAGSRLVWGTNGV 480

QY 481 VISIPLTETVWLRGQLLGLRANKTSPTSGEGARPGIITHVYDGDSSDRAASFIPYCSM 540
Db 481 VISIPLTETVWLRGQLLGLRANKTSPTSGEGARPGIITHVYDGDSSDRAASFIPYCSM 540

QY 541 AQAQLCFHGHRDAVKFFVSVPGNVLATLNGSVLDSPAEGPGPAAPASVEGQKLRNLVL 600
Db 541 AQAQLCFHGHRDAVKFFVSVPGNVLATLNGSVLDSPAEGPGPAAPASVEGQKLRNLVL 600

QY 601 SCGGGYIDFRICGDGEDDETERGAGDMSQVKPVLKSAERSHIIIVQVSYTPE 651
Db 601 SCGGGYIDFRICGDGEDDETERGAGDMSQVKPVLKSAERSHIIIVQVSYTPE 651

RESULT 2

AAAM41910
ID AAAM41910 standard; protein; 1066 AA.
XX
AC AAAM41910;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6841.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
WPI; 2001-442253/47.
DR N-PSDB; AAI61066.
XX
Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.
PT
XX
PS Example 2; SEQ ID NO 6841; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form

Qy 241 DPAPTPSGPQSGENGPEPDSSTRPEPESGDPGTGAGSAAPTMMLGAQNGWLYVHSA 300
Db 659 DPAPTPSGPQSGENGPEPDSSTRPEPESGDPGTGAGSAAPTMMLGAQNGWLYVHSA 718
Qy 301 VANWKKLHSLIKLSDVSLVHVKGRLVALADGTLAIFHRGEGQWDLNHYHLMDLGHP 360
Db 719 VANWKKLHSLIKLSDVSLVHVKGRLVALADGTLAIFHRGEGQWDLNHYHLMDLGHP 778
Qy 361 HHSIRCMVAVYDRVWCGYKKNVHVITQPKTMQIEKSFDAHPRESQVRQLAWIGDGVVYSI 420
Db 779 HHSIRCMVAVYDRVWCGYKKNVHVITQPKTMQIEKSFDAHPRESQVRQLAWIGDGVVYSI 838
Qy 421 RLDSTLRLYHAHTHOHLQDVIDIEPVVSKMLGTGKLGFSFVRITALLVAGSRILVWGTNGV 480
Db 839 RLDSTLRLYHAHTHOHLQDVIDIEPVVSKMLGTGKLGFSFVRITALLVAGSRILVWGTNGV 898
Qy 481 VISIPLTETVVLHRLGQLGLRANKTSPTSGEGARPGGIIHVYGDSSDRAASSFIPYCSM 540
Db 899 VISIPLTETVVLHRLGQLGLRANKTSPTSGEGARPGGIIHVYGDSSDRAASSFIPYCSM 958
Qy 541 AQAQLCFHGHDRDAVKFFVSPGNVLATLNGSVLDSPAEGPGPAAPASEVEGQKLRNLVL 600
Db 959 AQAQLCFHGHDRDAVKFFVSPGNVLATLNGSVLDSPAEGPGPAAPASEVEGQKLRNLVL 1018
Qy 601 SGGEGYIDFRIGDGEDDETEEGAGDMSQVKPVLKSAERSHIIWQVSYTPE 651
Db 1019 SGGEGYIDFRIGDGEDDETEEGAGDMSQVKPVLKSAERSHIIWQVSYTPE 1069

RESULT 4
ID AAM40124
XX AAM40124 standard; protein; 1014 AA.
AC AAM40124;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 3269.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PE 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI59280.

XX
PI Novel nucleic acids and polypeptides, useful for treating disorders such
PI as central nervous system injuries.
XX
PS Example 5; SEQ ID NO 3269; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1014 AA;
Query Match 99.8%; Score 3465; DB 4; Length 1014;
Best Local Similarity 99.8%; Pred. No. 3.4e-269;
Matches 650; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MKNVFPVYCRPLVEKOPTMKLWCAAGVNLGWRNEDDAGNGVKPAPGRDPLTCDREGD 60
Db 364 MKNVFPVYCRPLVEKOPTMKLWCAAGVNLGWRNEDDAGNGVKPAPGRDPLTCDREGD 423
Qy 61 GEPKSAHTSPEKKAKELPEMDATSSRWILTTSTLTSKVLIIDANQGTVVDDQFTVCA 120
Db 424 GEPKSAHTSPEKKAKELPEMDATSSRWILTTSTLTSKVLIIDANQGTVVDDQFTVCA 483
Qy 121 HVLCTISSIPAAASDSYPPGEMFLDSDVNPEPGADGVLGITVGCATRCNVPNSCSR 180
Db 484 HVLCTISSIPAAASDSYPPGEMFLDSDVNPEPGADGVLGITVGCATRCNVPNSCSR 543
Qy 181 GDTPLVDKGQEVATIANGKNVPSQSTTEATEATEVPDGPSEPETATLRPGPLEHVPT 240
Db 544 GDTPLVDKGQEVATIANGKNVPSQSTTEATEATEVPDGPSEPETATLRPGPLEHVPT 603
Qy 241 DPAPTPSGPQSGENGPEPDSSTRPEPESGDPGTGAGSAAPTMMLGAQNGWLYVHSA 300
Db 604 DPAPTPSGPQSGENGPEPDSSTRPEPESGDPGTGAGSAAPTMMLGAQNGWLYVHSA 663
Qy 301 VANWKKLHSLIKLSDVSLVHVKGRLVALADGTLAIFHRGEGQWDLNHYHLMDLGHP 360
Db 664 VANWKKLHSLIKLSDVSLVHVKGRLVALADGTLAIFHRGEGQWDLNHYHLMDLGHP 723
Qy 361 HHSIRCMVAVYDRVWCGYKKNVHVITQPKTMQIEKSFDAHPRESQVRQLAWIGDGVVYSI 420
Db 724 HHSIRCMVAVYDRVWCGYKKNVHVITQPKTMQIEKSFDAHPRESQVRQLAWIGDGVVYSI 783
Qy 421 RLDSTLRLYHAHTHOHLQDVIDIEPVVSKMLGTGKLGFSFVRITALLVAGSRILVWGTNGV 480
Db 784 RLDSTLRLYHAHTHOHLQDVIDIEPVVSKMLGTGKLGFSFVRITALLVAGSRILVWGTNGV 843
Qy 481 VISIPLTETVVLHRLGQLGLRANKTSPTSGEGARPGGIIHVYGDSSDRAASSFIPYCSM 540
Db 844 VISIPLTETVVLHRLGQLGLRANKTSPTSGEGARPGGIIHVYGDSSDRAASSFIPYCSM 903
Qy 541 AQAQLCFHGHDRDAVKFFVSPGNVLATLNGSVLDSPAEGPGPAAPASEVEGQKLRNLVL 600
Db 904 AQAQLCFHGHDRDAVKFFVSPGNVLATLNGSVLDSPAEGPGPAAPASEVEGQKLRNLVL 963
Qy 601 SGGEGYIDFRIGDGEDDETEEGAGDMSQVKPVLKSAERSHIIWQVSYTPE 651
Db 964 SGGEGYIDFRIGDGEDDETEEGAGDMSQVKPVLKSAERSHIIWQVSYTPE 1014

```

RESULT 5
ABG09904
ID ABG09904 standard; protein; 1069 AA.
XX
AC ABG09904;
XX
DT 13-FEB-2002 (first entry)
DE
DE Novel human diagnostic protein #9895.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
DR N-PSDB; AAS74091.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 40263; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1069 AA;

Query Match      97.0%; Score 3369; DB 4; Length 1069;
Best Local Similarity 97.7%; Pred. No. 1.9e-261;
Matches 636; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MKNVPVPCRPVVEKDPMTKLWCAAGVNLGWRPNEDDAGNVKPAFGPRDPLTCDREGD 60
Db 419 MKNVPVPCRPVSVGKDPMTLCLVAPANLNGWRPNEDDAGNVKPAFGPRDPLTSDREGD 478
Qy 61 GEPKSAHTSPKKAKELPENDATSSRWILTSLTTSKVVIIDANQFTVVDQFTVCNA 120
Db 479 GEPKSAHTSPKKAKELPENDATSSRWILTSLTTSKVVIIDANQFTVVDQFTVCNA 538

```

PT Polypeptides binding to Jun N-terminal protein kinase for treatment and
PT diagnosis of nervous system and inflammatory disorders.

XX Claim 21; Page 191-200; 254pp; Japanese.

XX The invention relates to polypeptides which bind to JNK3 (Jun N-terminal
XX protein kinase 3). The present sequence represents a JNK3 binding protein
XX (JSAP; JNK/SAPK associated protein) amino acid sequence. Included in the
XX invention are polypeptides derived from the JNK3 binding proteins, DNA
XX sequences encoding the proteins, expression vectors containing the DNA,
XX and host cells transformed by the vectors. The polypeptides exhibit
XX neurotropic, neuroprotective, antiparkinsonian, anticonvulsant,
XX cerebroprotective, neuroleptic, and antiinflammatory activity. JNK3
XX binding proteins are used in the treatment, prevention and diagnosis of
XX diseases associated with the JNK3 cascade, such as diseases of the
XX nervous system including Alzheimer's disease, Parkinson's disease,
XX Huntington's disease, multiple sclerosis, muscular dystrophy, apoplexy
XX and schizophrenia, dropsy, and inflammatory disorders

XX Sequence 1305 AA;

Query Match 91.9%; Score 3191; DB 3; Length 1305;
Best Local Similarity 91.2%; Pred. No. 5.1e-247;
Matches 594; Conservative 24; Mismatches 33; Indels 0; Gaps 0;

QY 1 MKNVVPVYCRPLVEKDPMTKLWCAAGVNLSCWRNEDDAGNGVPAQGRDPLTCDREGD 60
DB 655 MKNVVPVYCRPLVEKDPMTKLWCAAGVNLSCWRNEDDAGNGVPAQGRDPLTCDREGD 714

QY 61 GEPKSAHTSPEKKAKELPEMDATSRVWILTSTLTTSKVWIIDANQPGTVVDQFTVCNA 120
DB 715 GEPKSAHTSPEKKAKELPEMDATSRVWILTSTLTTSKVWIIDANQPGTVVDQFTVCNA 774

QY 121 HVLCTSSIPAAASDSYPPGEMFLSDVNPEDPGADGVLAGITLVGCATRCNVPKNCSSR 180
DB 775 HVLCTSSIPAAASDSYPPGEMFLSDVNPEDPGADGVLAGITLVGCATRCNVPKNCSSR 834

QY 181 GDTPLVDKGGQGVATTANGKVNPSQSTEEATEATEVPDPGPEPETATLRPGPLTEHVFT 240
DB 835 GDTPLVDKGGQGVATTANGKVNPSQSTEEATEATEVPDPGPEPETATLRPGPLTEHVFT 894

QY 241 DPAPTPSSGQPOGSENGPEPSSSTRPEPEPGDPTGAGSSAAPTMTLGAQNGWLYVHSA 300
DB 895 DPAPTPSSGQPOGSENGPEPSSSTRPEPEPGDPTGAGSSAAPTMTLGAQNGWLYVHSA 954

QY 301 VANWKKCLHSIKLQDSVLVHVGKRVILVALADGTILAFHREGDGQWDLNHYHMDLGH 360
DB 955 VANWKKCLHSIKLQDSVLVHVGKRVILVALADGTILAFHREGDGQWDLNHYHMDLGH 1014

QY 361 HHSIRCMVYDVRVWCGYKKNVHVTPKTMQIEKSFDAHPRESQVROLAWIGDGVVWSI 420
DB 1015 HHSIRCMVYDVRVWCGYKKNVHVTPKTMQIEKSFDAHPRESQVROLAWIGDGVVWSI 1074

QY 421 RLDSTLRLYHAHTHQLQDVIDIEPVYSKMLGTCKLGFSEVRITALLVAGSRLWVGTVNGV 480
DB 1075 RLDSTLRLYHAHTHQLQDVIDIEPVYSKMLGTCKLGFSEVRITALLVAGSRLWVGTVNGV 1134

QY 481 VTSIPLTETVVLHRRGQLGLRANKTSPTSGEGARPGGIIHVYGDSDSDRAASFPYCSM 540
DB 1135 VTSIPLTETVVLHRRGQLGLRANKTSPTSGEGARPGGIIHVYGDSDSDRAASFPYCSM 1194

QY 541 AQAQCLCFHGRDAVAFVSPGNVLATNGSVLSDSPAEGGPAAPASEVEGOKLRNLVL 600
DB 1195 AQAQCLCFHGRDAVAFVSPGNVLATNGSVLSDSPAEGGPAAPASEVEGOKLRNLVL 1254

QY 601 SGSEGYIDFRIGDEDETEEGAGDMSQVKPVLKSAERSHIIWVQVSYTPE 651
DB 1255 SGSEGYIDFRIGDEDETEEGAGDMSQVKPVLKSAERSHIIWVQVSYTPE 1305

RESULT 7
AAB12876
ID AAB12876 standard; protein; 1314 AA.

XX AAB12876;

XX 01-NOV-2000 (first entry)

XX Murine JNK3 binding protein amino acid sequence #2.

XX JNK3; Jun N-terminal protein kinase 3; JNK3 binding protein; JSAP;
XX JNK/SAPK associated protein; neurotropic; neuroprotective; neuroleptic;
XX antiparkinsonian; anticonvulsant; cerebroprotective; antiinflammatory;
XX nervous system disease; Alzheimer's disease; Parkinson's disease;
XX Huntington's disease; multiple sclerosis; muscular dystrophy; apoplexy;
XX schizophrenia; dropsy; inflammatory disorders; mouse.

XX Mus sp.

XX WO200031132-A1.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-JP006487.

XX 24-NOV-1998; 98JP-00332484.

XX 02-SEP-1999; 99JP-00248442.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Ichimura M, Hirose R, Yoshioka K;

XX WPI; 2000-400042/34.

XX N-PSDB; AAA62900.

XX Polypeptides binding to Jun N-terminal protein kinase for treatment and
XX diagnosis of nervous system and inflammatory disorders.

XX Claim 1; Page 200-210; 254pp; Japanese.

XX The invention relates to polypeptides which bind to JNK3 (Jun N-terminal
XX protein kinase 3). The present sequence represents a JNK3 binding protein
XX (JSAP; JNK/SAPK associated protein) amino acid sequence. Included in the
XX invention are polypeptides derived from the JNK3 binding proteins, DNA
XX sequences encoding the proteins, expression vectors containing the DNA,
XX and host cells transformed by the vectors. The polypeptides exhibit
XX neurotropic, neuroprotective, antiparkinsonian, anticonvulsant,
XX cerebroprotective, neuroleptic, and antiinflammatory activity. JNK3
XX binding proteins are used in the treatment, prevention and diagnosis of
XX diseases associated with the JNK3 cascade, such as diseases of the
XX nervous system including Alzheimer's disease, Parkinson's disease,
XX Huntington's disease, multiple sclerosis, muscular dystrophy, apoplexy
XX and schizophrenia, dropsy, and inflammatory disorders

XX Sequence 1314 AA;

Query Match 91.9%; Score 3191; DB 3; Length 1314;
Best Local Similarity 91.2%; Pred. No. 5.2e-247;
Matches 594; Conservative 24; Mismatches 33; Indels 0; Gaps 0;

QY 1 MKNVVPVYCRPLVEKDPMTKLWCAAGVNLSCWRNEDDAGNGVPAQGRDPLTCDREGD 60
DB 664 MKNVVPVYCRPLVEKDPMTKLWCAAGVNLSCWRNEDDAGNGVPAQGRDPLTCDREGD 723

QY 61 GEPKSAHTSPEKKAKELPEMDATSRVWILTSTLTTSKVWIIDANQPGTVVDQFTVCNA 120
DB 724 GEPKSAHTSPEKKAKELPEMDATSRVWILTSTLTTSKVWIIDANQPGTVVDQFTVCNA 783

QY 121 HVLCTSSIPAAASDSYPPGEMFLSDVNPEDPGADGVLAGITLVGCATRCNVPKNCSSR 180
DB 784 HVLCTSSIPAAASDSYPPGEMFLSDVNPEDPGADGVLAGITLVGCATRCNVPKNCSSR 843

QY 181 GDTPLVDKGGQGVATTANGKVNPSQSTEEATEATEVPDPGPEPETATLRPGPLTEHVFT 240
DB 844 GDTPLVDKGGQGVATTANGKVNPSQSTEEATEATEVPDPGPEPETATLRPGPLTEHVFT 903

KW antiparkinsonian; anticonvulsant; cerebroprotective; antiinflammatory;
KW nervous system disease; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; multiple sclerosis; muscular dystrophy; apoplexy;
KW schizophrenia; dropsy; inflammatory disorders; mouse.
XX Mus sp.
XX WO200031132-A1.
XX 02-JUN-2000.
XX 19-NOV-1999; 99WO-JP006487.
XX 24-NOV-1998; 98JP-00332484.
XX 02-SEP-1999; 99JP-00248442.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Ichimura M, Hirose R, Yoshioka K;
XX WPI; 2000-400042/34.
XX N-PSDB; AAA62901.
XX Polypeptides binding to Jun N-terminal protein kinase for treatment and
XX diagnosis of nervous system and inflammatory disorders.
XX Claim 1; Page 210-220; 254pp; Japanese.
XX The invention relates to polypeptides which bind to JNK3 (Jun N-terminal
XX protein kinase 3). The present sequence represents a JNK3 binding protein
XX (JSAP; JNK/SAPK associated protein) amino acid sequence. Included in the
XX invention are polypeptides derived from the JNK3 binding proteins, DNA
XX sequences encoding the proteins, expression vectors containing the DNA,
XX and host cells transformed by the vectors. The polypeptides exhibit
XX neurotropic, neuroprotective, antiparkinsonian, anticonvulsant,
XX cerebroprotective, neuroleptic, and antiinflammatory activity. JNK3
XX binding proteins are used in the treatment, prevention and diagnosis of
XX diseases associated with the JNK3 cascade, such as diseases of the
XX nervous system including Alzheimer's disease, Parkinson's disease,
XX Huntington's disease, multiple sclerosis, muscular dystrophy, apoplexy
XX and schizophrenia, dropsy, and inflammatory disorders
XX Sequence 1337 AA;
Query Match 91.9%; Score 3191; DB 3; Length 1337;
Best Local Similarity 91.2%; Pred. No. 5.3e-247;
Matches 594; Conservative 24; Mismatches 33; Indels 0; Gaps 0;
QY 1 MKNVPVYVYCRPLVEKDPKMLKCAAGVNLGWRNEDDAGNGVKPAPGRDPLTCDREGD 60
DB 687 MKNVPVYVYCRPLVEKDPKMLKCAAGVNLGWRNEDDAGNGVKPAPGRDPLTCDREGD 746
QY 61 GEPKSAHTSPKPKKAKELPEMDATSSRVWILSTLTTSKVIIDANQGTVDVQFTVCNA 120
DB 747 GEPKSTHPSPEKKAKETPEADATSSRVWILSTLTTSKVIIDANQGTVDVQFTVCNA 806
QY 121 HVLCTSIIPAAASDSDYPPGEMFLDSVNPEDGADGVLGITLVGCATCNVPRNSCSR 180
DB 807 HVLCTSIIPAAASDSDYPPGEMFLDSVNPEDGADGVLGITLVGCATCNVPRNSCSR 866
QY 181 GTPVLDKQGEVATLANGKWPNSQSTEBATEATEVPDPGSEPEATLTPGPLETHVFT 240
DB 867 GTPVLDKQGEVATLANGKWPNSQSTEBATEATEVPDPGSEPEATLTPGPLETHVFT 926
QY 241 DPAPTSSGPQSGENGPDPSSSTRPEPEPSGDPGTGAGSSAAPTWLGAQNGWLYVHSA 300
DB 927 DPAPTSSSTQPAESNGESNGTIVQVPEPSGELSTTTSSAAPTWMLGAQNGWLYVHSA 986
QY 301 VANWKKCLHSIKLSDVLSLVHVHVKGRVLVALADGTLAIPIHRGSDGWDLSNVHLMDLGHP 360
DB 987 VANWKKCLHSIKLSDVLSLVHVHVKGRVLVALADGTLAIPIHRGSDGWDLSNVHLMDLGHP 1046
QY 361 HHSIRCMVYVDRVWCGYKKNKHVHVIQPKTMQIEKSFDAHPRESQVRLAWIGDGVWVSI 420

Db 1047 HHSIRCMVYVDRVWCGYKKNKHVHVIQPKTMQIEKSFDAHPRESQVRLAWIGDGVWVSI 1106
QY 421 RLDSTLRLYHATHQHLOVDIEPYVSKMLGTGKLGFSFVRITALLVAGSLWGTGNGV 480
Db 1107 RLDSTLRLYHATHQHLOVDIEPYVSKMLGTGKLGFSFVRITALLVAGSLWGTGNGV 1166
QY 481 VISIPLTETVVLHRLGQLLGRANKTSPTSRGARPGGIHVVYDDSSDRAASSFTPYCSM 540
Db 1167 VISIPLTETVVLHRLGQLLGRANKTSPTSRGARPGGIHVVYDDSSDRAASSFTPYCSM 1226
QY 541 AAQQLCFHGHDRDAVKFFVSPGNVLTALNGSVLPSPAECPGAPAPASVEGOKURNVLVL 600
Db 1227 AAQQLCFHGHDRDAVKFFVSPGNVLTALNGSVLPSPAECPGAPAPASVEGOKURNVLVL 1286
QY 601 SGGEGYIDFRIGDGEDDETEGAGDMSQVKVLSKAERSHIIIVQVSYTPE 651
Db 1287 SGGEGYIDFRIGDGEDDETEGAGDMSQVKVLSKAERSHIIIVQVSYTPE 1337
RESULT 10
ABG09902
ID ABG09902 standard; protein; 809 AA.
XX AC
XX ABG09902;
XX DT 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #9893.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS74089.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 20; SEQ ID NO 40261; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity

CC ADA45192). JLP tethers MEK kinase 3 (MEKK3), Mitogen-Activated Protein
CC kinase (MAPK), c-Jun NH2-terminal kinase (JNK), p38 MAP
CC kinase (MAPK), c-Myc and MAX into a signalling module which controls the
CC apoptotic response. JLP therefore functions as a signalling conduit to
CC transmit extracellular signals to the nucleus through MEKK3-MKK4-
CC JNK/p38/MAPK/c-Myc/MAX signalling module. The JLP sequences are useful
CC for modulating apoptotic response in a cell, and thus for treating
CC metastatic cancer. The present sequence is a C-terminal deletion mutant
CC of murine JLP, generated in an example from the invention.
XX
SQ Sequence 1165 AA;

Query Match 42.2%; Score 1467; DB 6; Length 1165;
Best Local Similarity 57.3%; Pred. No. 1.6e-108; Indels 44; Gaps 12;
Matches 305; Conservative 68; Mismatches 115;
QY 1 MKNVPVYVCRPLVEKPTMKLWCAAGVNLGWRPNEDDAGNGVKPAP-GRDPLTCDREG 59
DB 652 MKNLPVYVLRPLDEKDTSMKLCWCAVGNLGGKTR--DGGSVVGASVFKDIAGLDTEG 709
QY 60 DGEPKSAHTSP-----KKAKELPEMDATSSRWILTSILTTSKVLIIDANQPGTVVD 113
DB 710 SKQRSASQSLDKLDQELKEQKEPKNQEELSSQVWICTSTHTTKVIIIDAVQPGNILD 769
QY 114 QFTVCNAHVLCISSTIPAAASDSYPPGEMFLDSD-----VNPEDPGADGVLGIT 162
DB 770 SFTVCNSHVLCIASVPGARETDYPAGEELSSQVWIKASLCSGSMTSNSAEWDSLGGIT 829
QY 163 LVGCATRCNVPNSCRRGDTPLVDKQGEVATTANGKVNPSQST-EEATEATEVPPDG 220
DB 830 VVGCSSTGLTGAATSPSTNGASPVIEK-PPMET-ENSEVDENIPTAEATEATE-GNAG 886
QY 221 PSEPTATLRPGPLTEHVFTDPAPTP-----SSGPOGSENGPEPDSSSTRP-EPEPSGD 274
DB 887 STEDTVDISQGVYTEHVFTDPLGVQIPEDLSPVQSSNDSVYKQDQISVLNEQDLARE 946
QY 275 PTGASSAAPTMLGAQNGWLVHSAVANWKKLHSLIKLSDSVLSVHVKGRLVALADG 334
DB 947 EAQKMSLLPTMLGAQNGCLVHSSVAQWRKCLHSIKLSDLSIVHVKGIVLVALADG 1006
QY 335 TLAIFHRGEDGQWDLNHYLMDLGHPPHHSIRCMAYVYDRVWCGYKKNVHVIOPTKMQIEK 394
DB 1007 TLAIFHRGVDGQWDLNHYLMDLGRPHHSIRCMVTVVHDKVWCGYKKNVHVIOPTKMKIEK 1066
QY 395 SFDAPHRQESQVROLAWIGDGVVWVIRLDSTLRLYHAHTHOHLQDVIDEPPYVSKMLGTGK 454
DB 1067 SFDAPHRKESQVROLAWIGDGVVWVIRLDSTLRLYHAHTYQHLQDVIDEPPYVSKMLGTGK 1126
QY 455 LGFSFVRITALLVAGSRLWVGTVGNGVWISIPLETETVWLHRLGQLGLRANKTS 506
DB 1127 LGFSFVRITALLVAGSRLWVGTVGNGVWISIPLETET-----NKTS 1165

RESULT 15
AAW64473
ID AAW64473 standard; protein; 828 AA.
XX
AC AAW64473;
XX
DT 16-OCT-1998 (first entry)
XX
DE Human secreted protein from clone EC172_1.
XX
KW Secreted protein; nutrition; cytokine; cell proliferation; activin;
KW differentiation; immune system; stimulator; suppressor; tissue growth;
KW haematopoiesis regulation; inhibin; chemotactic; chemokinetic; cadherin;
KW haemostatic; thrombolysis; receptor; ligand; anti-inflammatory.
XX
OS Homo sapiens.
XX
PN WO9830589-A2.
XX
PD 16-JUL-1998.

XX 20-DEC-1997; 97WO-US023506.
XX
XX 10-JAN-1997; 97US-00781225.
PR 18-DEC-1997; 97US-00993228.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
PI Spaulding V, Agostino MJ;
PI
PI WPI; 1998-413682/35.
DR N-PSDB; AAV46322.
XX
XX New isolated nucleic acids and secreted proteins - isolated from human
PT foetal kidney, adult placenta, adult colon, adult testes, foetal brain
PT and adult brain cDNA libraries.
XX
XX Claim 39a; Page 97-100; 122pp; English.
XX
XX This sequence represents a novel secreted protein from clone EC172_1
CC isolated from a human adult brain cDNA library. This novel protein may
CC have biological activity, e.g. nutritional activity, cytokine and cell
CC proliferation/differentiation activity, immune stimulating or suppressing
CC activity, haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity or other activities
XX
SQ Sequence 828 AA;

Query Match 41.6%; Score 1446.5; DB 2; Length 828;
Best Local Similarity 57.7%; Pred. No. 4.5e-107;
Matches 300; Conservative 65; Mismatches 124; Indels 31; Gaps 10;
QY 1 MKNVPVYVCRPLVEKPTMKLWCAAGVNLGWRPNEDDAGNGVKPAP-GRDPLTCDREG 59
DB 314 MKNLPVYVLRPLDEKDTSMKLCWCAVGNLGGKTR--DGGSVVGASVFKDVAGLDTEG 371
QY 60 DGEPKSAHTSP-----KKAKELPEMDATSSRWILTSILTTSKVLIIDANQPGTVVD 113
DB 372 SKQRSASQSLDKLDQELKEQKEPKNQEELSSQVWICTSTHTSAATKVIIDAVQPGNILD 431
QY 114 QFTVCNAHVLCISSTIPAAASDSYPPGEMFLDSD-----VNPEDPGADGVLGIT 162
DB 432 SFTVCNSHVLCIASVPGARETDYPAGEELSSQVWIKASLCSGSMTSNSAEWDSLGGIT 491
QY 163 LVGC-ATRCNVPNSCRRGDTPLVDKQGEVATTANGKVNPSQST-EEATEATEVPPDG 220
DB 492 VVGCSAEGVTGAATSPSTNGASPVMDKPPMEA--ENSEVDENIPTAEATEATE-GNAG 548
QY 221 PSEPTATLRPGPLTEHVFTDPAPTP-----SSGPOGSENGPEPDSSSTRP-EPEPSGD 274
DB 549 SAEDTVDISQGVYTEHVFTDPLGVQIPEDLSPVQSSNDSVYKQDQISVLNEQDLARE 608
QY 275 PTGASSAAPTMLGAQNGWLVHSAVANWKKLHSLIKLSDSVLSVHVKGRLVALADG 334
DB 609 EAQKMSLLPTMLGAQNGCLVHSSVAQWRKCLHSIKLSDLSIVHVKGIVLVALADG 668
QY 335 TLAIFHRGEDGQWDLNHYLMDLGHPPHHSIRCMAYVYDRVWCGYKKNVHVIOPTKMQIEK 394
DB 669 TLAIFHRGVDGQWDLNHYLMDLGRPHHSIRCMVTVVHDKVWCGYKKNVHVIOPTKMKIEK 728
QY 395 SFDAPHRQESQVROLAWIGDGVVWVIRLDSTLRLYHAHTHOHLQDVIDEPPYVSKMLGTGK 454
DB 729 SFDAPHRKESQVROLAWIGDGVVWVIRLDSTLRLYHAHTYQHLQDVIDEPPYVSKMLGTGK 788
QY 455 LGFSFVRITALLVAGSRLWVGTVGNGVWISIPLETETVWLHRL 494
DB 789 LGFSFVRITALLVAGSRLWVGTVGNGVWISIPLETESKIFR 828

Search completed: August 23, 2004, 11:11:35
Job time : 65 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 11:11:38 ; Search time 55 seconds
(without alignments)
3719.637 Million cell updates/sec

Title: US-10-019-495-9

Perfect score: 3473

Sequence: 1 MKNVPVPPVYCRPLVEKPTM.....VLSKAERSHIIWQVSYTPE 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	1732.5	49.9	1277	12	US-10-276-774-2154
2	807.5	23.3	235	9	US-09-925-297-626
3	511	14.7	118	15	US-10-264-237-2212
4	460.5	13.3	766	14	US-10-205-219-157
5	442.5	12.7	2063	12	US-09-918-715-204
6	333.5	9.6	1240	12	US-10-114-270-62
7	333.5	9.6	1284	16	US-10-408-765A-1008
8	305.5	8.8	1121	16	US-10-408-765A-865
9	304.5	8.8	1119	16	US-10-408-765A-2438
10	186.5	5.4	119	15	US-10-264-237-1721
11	185.5	5.3	52	9	US-09-864-761-44323
12	161	4.6	1809	12	US-10-114-270-98
13	161	4.6	1829	12	US-10-312-352-16
14	154.5	4.4	1228	10	US-09-917-384-1
15	154.5	4.4	1228	10	US-09-917-383-1

16	148.5	4.3	1740	14	US-10-192-381-40	Sequence 40, Appl
17	146.5	4.2	985	10	US-09-994-064-6	Sequence 6, Appl
18	146.5	4.2	985	10	US-09-994-064-66	Sequence 66, Appl
19	142	4.1	3551	12	US-10-263-929-144	Sequence 144, App
20	141.5	4.1	582	14	US-10-156-761-11938	Sequence 11938, A
21	140	4.0	668	16	US-10-437-963-134726	Sequence 134726, A
22	138	4.0	1367	9	US-09-801-368-108	Sequence 108, App
23	138	4.0	1538	16	US-10-437-963-144952	Sequence 144952, A
24	135.5	3.9	406	7	US-10-264-237-2046	Sequence 2046, Ap
25	134	3.9	406	8	US-08-523-004-2	Sequence 2, Appli
26	134	3.9	406	8	US-08-875-849C-2	Sequence 2, Appli
27	134	3.9	406	9	US-09-859-214-68	Sequence 68, Appl
28	134	3.9	406	10	US-09-862-989-2	Sequence 2, Appli
29	134	3.9	406	13	US-10-047-542-68	Sequence 68, Appl
30	134	3.9	965	16	US-10-437-963-190305	Sequence 190305, A
31	133.5	3.8	625	16	US-10-437-963-165015	Sequence 165015, A
32	133	3.8	382	7	US-08-523-004-4	Sequence 4, Appli
33	133	3.8	382	8	US-08-875-849C-4	Sequence 4, Appli
34	133	3.8	382	9	US-09-859-214-70	Sequence 70, Appl
35	133	3.8	382	10	US-09-862-989-4	Sequence 4, Appli
36	133	3.8	731	13	US-10-086-464-8	Sequence 8, Appli
37	132.5	3.8	381	14	US-10-165-603-20	Sequence 20, Appl
38	132.5	3.8	1747	16	US-10-437-963-168997	Sequence 168997, A
39	132	3.8	591	12	US-10-425-114-68271	Sequence 68271, A
40	132	3.8	1151	12	US-09-825-751A-79	Sequence 79, Appl
41	131.5	3.8	1555	16	US-10-437-963-129858	Sequence 129858, A
42	131.5	3.8	1974	16	US-10-647-196-24	Sequence 24, Appl
43	131	3.8	584	14	US-10-156-761-12405	Sequence 12405, A
44	131	3.8	1843	16	US-10-437-963-194983	Sequence 194983, A
45	130.5	3.8	297	16	US-10-437-963-168840	Sequence 168840, A

ALIGNMENTS

RESULT 1

US-10-276-774-2154
; Sequence 2154, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2154
; LENGTH: 1277
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2154

Query Match	49.9%	Score 1732.5	DB 12	Length 1277
Best Local Similarity	57.3%	Pred. No. 11e-127		
Matches	358	Conservative	79	Mismatches 141
				Indels 47
				Gaps 13
QY	1	MKNVPVPPVYCRPLVEKPTM	1	MKNVPVPPVYCRPLVEKPTM
Db	664	MKNLPVPLRPLDEKDTSMKLCVAGVNLGGKTR	1	DGGSVVGASVFFYKDVAGLDTG 721
QY	60	DGEFKSAHTSPE	1	KKAKELPMDATSRVWLTSTLTTSKVIIDANQPTVVD 113
Db	722	SKORASQSSLDKLDQELKEQKELKNQBELSVWCTSTHSAKVLIIIDAVQPGNILD 781		
QY	114	QFTVCNAHVLCISIPAAASDYPGEMFLDSD	1	VNPEDFGADVLAGIT 162
Db	782	SFIVCSHVLCIASVFGARETDIPAGEDLSGQVOKASLCGSMTNSNSAETSLGGIT 841		


```

; FEATURE:
; OTHER INFORMATION: Sperm-specific protein
US-10-205-219-157

Query Match      13.3%; Score 460.5; DB 14; Length 766;
Best Local Similarity 42.7%; Pred. No. 2.4e-27;
Matches 112; Conservative 40; Mismatches 85; Indels 25; Gaps 8;

QY 1 MKNVPPVYCRPLVEKDPMTKMLWCAAGVNLGWRPNEDDAGNGVKAP--GRDPLTCDREG 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 487 MKNLFPVYLRDLKDKTSKMLWCAVGVLGSKTR--DGSVVGASVPYKDVAGLDREG 544
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 60 DGEPSAHTSPE-----KKKAKELPEMDATSRVWILSTLTTSKVVIIDANQPGTVVD 113
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 545 SKORGASQSLDKLQELKEQCKELKNQBELSSLVWICTSTHSATKVLIIIDAVQGNILD 604
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 114 QFTVCNAHVLCSSIPAAASDSYPPGCEMFLDS-----VNPDGADGVLAGIT 162
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 605 SFTVCNSHVLCIASPGARETDYPAGEDLSGQVDKASLCGSMTSNSAETDSLGGIT 664
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 163 LVGC-ATRCNVPRNSCSRGDTPVLDDKQGEVATIANGKVNPSQST--PEATEATEVPDPG 220
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 665 VVGCASAGVTGAATSPSTNGASVMDKPPMEA--ENSEVDENVPTAEATEATE--GNAG 721

QY 221 PSEPETALRPGPLTEHVFTDP 242
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 722 SAEDTVDISQTVGTVTEHVFTDP 743
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
US-09-918-715-204
; Sequence 204, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 2063
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-715-204

Query Match      12.7%; Score 442.5; DB 12; Length 2063;
Best Local Similarity 26.4%; Pred. No. 2.6e-25;
Matches 168; Conservative 86; Mismatches 223; Indels 159; Gaps 25;

QY 85 SSRVWILSTLTTSKVVIIDANQPGTVVDQFTVCNAHVLCISSIPAAASDSY--PPGSMF 142
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1513 SPEVWVCNSDGYVGQVCLLSRAEPDVEACIACVCSARILCIGAVPGLQPRCHREPPSLR 1572
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 143 LDSVNPEDPGADGVLAGITLVGCATRCNVPRNSC-----SSRGDTPVLDDK----- 189
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1573 SPPETAPEAGPELDVEAAADEEAATLAPGPQPCLIHSIAGSLGEMTFGLGEGDPRPEL 1632

QY 190 -----QGEVATIANGKVN---PS---QSTEEATEATEVPDPGPSE 223
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1633 VPFDSDSDDESSPSSGTLOQASRSTISSFGNEETPSSKEATATSTSSSEEQEPG--FL 1691
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 224 PETATLRPGPLTEHVFTDPAPTSSGPGQSGENGPEPDSSSTRPEPEP--SGDPTGAGSS 281
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Db 1692 PLSGSFGPG-----GPCGT---SPMDGRALRRSSHSGSFTRGSLDILLVDPPEAYQSS 1740
QY 282 AAPTWLGQONGLVYVHSAVANWKKLHSIKLKD--SVLSLVHVKGRVLVALADGTALIF 339
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1741 ----VMLGTEDGCVHYQSSDSIRDRRNSMKLQHAASVTCTILYLNQVFSVLANGELVYV 1796
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 340 HRGEDQMDLSNYHMLDGLGHPHHSIRCMVAVYDRVMCGYKNKVHVITQPKTMQIEKSFDAH 399
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1797 QREAGHFMDPQFKSVTLTGTSQSPITKMWVSVGGRLWCGCQNRVLVLSPTTLQLEHFYV- 1855
QY 400 PRRESQVROLAWIGD---GVVVSIRLDTLRLYHARTHQHLQDQDIEPVVSKML--GTGKL 455
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1856 --GQDSSRCVACWMDSSSLGVWVTLKGSAHVCLYHPDTEQLAEVDVTPPVHRLAGSDAI 1913
QY 456 ----GFSFVRITALLVAGSRILWGTGNGVVISIPITE--TVLHRRGOLLGLRANKTSPTSG 510
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1914 IRQKAAACILRIITALLVCELLWVGTSGVLTMTPTSPGTVCPRAPL-----SPT-G 1964
QY 511 EGARPGGIHVYGDSDSDRAASSFTPYCSMAQAQLCFCHGRDAVKFFVSV----- 560
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1965 LG-----QGHTHGVRFLLAAVQLPDGPNLLC 1989
QY 561 ----PGNV---LATNGSVLDSF--ABGPGPAPAPASEVEGQKLRNVLVLSGEGVIDPR 610
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1990 PTPPPPPDTPGPKLPSLEHR--DSPWHRGPAFAP------KMLVISGSDGYEDPR 2037
QY 611 IDGEDDDETEGAGDMQSVKPVLSKAERSHIIVQV 646
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2038 LSSGGSSSETVGRDSDS-----TNHLLMRV 2063

RESULT 6
US-10-114-270-62
; Sequence 62, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Zhaozhong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zethusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liete, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
```

```
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 62
; LENGTH: 1240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-62
```

```
Query Match          9.6%; Score 333.5; DB 12; Length 1240;
Best Local Similarity 24.7%; Pred. No. 5.4e-17;
Matches 120; Conservative 72; Mismatches 169; Indels 125; Gaps 17;

QY 230 RPSPTVKSF-----PLAAPVLCMEYIPELEEEAESRDESVADPS---ATVPTICL 288
Db 807 RPSPTVKSF-----PLAAPVLCMEYIPELEEEAESRDESVADPS---ATVPTICL 857
QY 289 GAQGNLXVHSAVANWKKLHSLK--DSVLSLVHVGRVULVALADGTLAIFHRGEGD- 345
Db 858 GLQDSILLYSSVDTGTQCLVSCRSPGLQVLCRLHSPHLLAGLQDGTAAAYPRTSGV 917
QY 346 QWDLNLYHMLDGLHPHHSIRCMVYVDRVWCYKKNVHVIOPTKMQIEKSPDAHRRSQ 405
Db 918 LWDLESPVCLTVGP-GPVRTLLSLEDVAVWASCGPRVTVEATTLQPOQSFHQADEAVS 976
QY 406 VRQLAWIGDGVVSVIRLDTLRLYHAHTHQHDVDIEPVYVKMLGTGKLGFSFVRITAL 465
Db 977 VTHVMKAGVWMAFSSGTSIRLFTETLEHLQEIATRTTFLP-----GQKHLCVTSL 1032
QY 466 LVAGSRLWVGTVGNVVISIPLTETVVL-----HRG-----QLLGLRAN 503
Db 1033 LICQGLLWGTQGVVLLPVRLEGIPKITGKGMVSLNGHCGPVAFLAVATSILAPDIL 1092
QY 504 KTSPTSGEGAR-----PGGIHVYGDSS-----DRAASSFIPYCSMAQAOLCFHGRDA 553
Db 1093 RSDQEAEGPRAEEDKPDQQAHPMPDHSVHGRELTRKKGILLQYLRSTAHL-----1144
QY 554 VKFFSVFPGNVLATLNGSVLDSPAEGPAPAAASEVEG-----591
Db 1145 -----FQPLLS-----MRPAPADGAALHSESDGSIEYEMADDPDWWVRSRPCARDA 1191
QY 592 --QKURNVLVSGGEGYIDF--RIGD-----GEDDETEEGAGDMSQVKPVLKASRSH 640
Db 1192 HRKEICSAIISGGGYRNFSGALSGRQAPCGTDTST-----1230
QY 641 IIVWQV 646
Db 1231 LLIWQV 1236
```

```
RESULT 7
US-10-408-765A-1008
; Sequence 1008, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
```

```
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1008
; LENGTH: 1284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1008
```

```
Query Match          9.6%; Score 333.5; DB 16; Length 1284;
Best Local Similarity 24.7%; Pred. No. 5.7e-17;
Matches 120; Conservative 72; Mismatches 169; Indels 125; Gaps 17;

QY 230 RPSPTVKSF-----PLAAPVLCMEYIPELEEEAESRDESVADPS---ATVPTICL 288
Db 851 RPSPTVKSF-----PLAAPVLCMEYIPELEEEAESRDESVADPS---ATVPTICL 901
QY 289 GAQGNLXVHSAVANWKKLHSLK--DSVLSLVHVGRVULVALADGTLAIFHRGEGD- 345
Db 902 GLQDSILLYSSVDTGTQCLVSCRSPGLQVLCRLHSPHLLAGLQDGTAAAYPRTSGV 961
QY 346 QWDLNLYHMLDGLHPHHSIRCMVYVDRVWCYKKNVHVIOPTKMQIEKSPDAHRRSQ 405
Db 962 LWDLESPVCLTVGP-GPVRTLLSLEDVAVWASCGPRVTVEATTLQPOQSFHQADEAVS 1020
QY 406 VRQLAWIGDGVVSVIRLDTLRLYHAHTHQHDVDIEPVYVKMLGTGKLGFSFVRITAL 465
Db 1021 VTHVMKAGVWMAFSSGTSIRLFTETLEHLQEIATRTTFLP-----GQKHLCVTSL 1076
QY 466 LVAGSRLWVGTVGNVVISIPLTETVVL-----HRG-----QLLGLRAN 503
Db 1077 LICQGLLWGTQGVVLLPVRLEGIPKITGKGMVSLNGHCGPVAFLAVATSILAPDIL 1136
QY 504 KTSPTSGEGAR-----PGGIHVYGDSS-----DRAASSFIPYCSMAQAOLCFHGRDA 553
Db 1137 RSDQEAEGPRAEEDKPDQQAHPMPDHSVHGRELTRKKGILLQYLRSTAHL-----1188
QY 554 VKFFSVFPGNVLATLNGSVLDSPAEGPAPAAASEVEG-----591
Db 1189 -----FQPLLS-----MRPAPADGAALHSESDGSIEYEMADDPDWWVRSRPCARDA 1235
QY 592 --QKURNVLVSGGEGYIDF--RIGD-----GEDDETEEGAGDMSQVKPVLKASRSH 640
Db 1236 HRKEICSAIISGGGYRNFSGALSGRQAPCGTDTST-----1274
QY 641 IIVWQV 646
Db 1275 LLIWQV 1280
```

```
RESULT 8
US-10-408-765A-865
; Sequence 865, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
```

; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 865
 ; LENGTH: 1121
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-865

Query Match
 Best Local Similarity 8.8%; Score 305.5; DB 16; Length 1121;
 Matches 134; Conservative 81; Mismatches 192; Indels 227; Gaps 22;

QY	29	NLSGRWRNPDAGNVKAP-----GRDPLTCDEGDEPKSAHTSPKPKAKE-LPEMD	82
Db	610	NHMGWFCVEDD-GNHIKKEKHPLLVGHMPVMAKQOEKIECAAYNPFPYLNNESQPSDF	668
QY	83	ATS-SRWILSTLTTSKVIID-ANQPGTVVDQTCNAHVLCISSIPAAASDSYPFGE	140
Db	669	STAHGFLWIGSCTHQMGOIAIVSFQNSTPKVIECFNV-ESRILCMLYVPV-----	717
QY	141	MFLDSVNPDPGADGVLAGITLVGCATRCNVPNSCSRSGDTPVLDKGQGEVATIANGK	200
Db	718	-----EKKRREPGA-----	726
QY	201	VNPQSSTEEATEVDPGPEPETATLRPGPLTEHVFTDPATPSSGPQGSNGPEP	260
Db	727	-----PPDETFAVR-----	736
QY	261	DSSSTRPEPEPGDPTGAGSAAPTMMLGAQNGMYVHSVANWKKCLHSIKL-----	313
Db	737	-----ASDVPTICVTEGSIYKSSQGSKK-----VRLQHFPTPE	773
QY	314	KDSVLSLVHVKGRVLVALADGTALFHRGEGDQWDLNHYHMDLGHPPHHSIRCMVAVYDR	373
Db	774	KSTVMSLACTSQSLYAGLVNGAVASYARAPDGSMDSEPKVKILG--VLPVRSLLMMEDT	831
QY	374	VMCGYKKNKHVIOPTQWIEKSFDAHPRESQVQLAWIGDGVWVSRILDSLTIRLYHAHT	433
Db	832	LWAASGGQVFIISVETHAVEGLEAHQEBGVISHMAVSGVGIWIAFTSGSTLRLFTET	891
QY	434	HQHLQDVVDIEPYVKMLGTGKLGFSFVRITALLVAGSRWVGTGNGVVISIPLTETVLH	493
Db	892	LKHLQDINIATPVHMLP-----GHQRLSVTSLLVCHGLLVGTSLGVLVLPV-----	940
QY	494	RGQLLGRANKTSPTSSEGAPPGIIHVYGDSSDRAASSFIPYCSMAQAQLCFHGRDA	553
Db	941	-----PRLOGIPKVTG-----	963
QY	554	VKFPVSVPGNVLATL-----NGSVLDSPAEGPGPAAPASEVEGOKLRNVLSGEGYID	608
Db	964	VKFI-----VLATALHEKDKKSRDSLAPGPEP-----QDEDQK-----DALPSGGAG---	1006
QY	609	FRIGDGEDDET---EAGAGDMSQVKPVLSKAERS	639
Db	1007	SSLSQGDPAAIWLGDSIGSMTQ-KSDLSSSSGS	1039

RESULT 9

US-10-408-765A-2438
 ; Sequence 2438, Application US/10408765A
 ; Publication No. US20040101874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Eoin D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2438
 ; LENGTH: 1121
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-2438

Query Match
 Best Local Similarity 9.8%; Score 304.5; DB 16; Length 1121;
 Matches 134; Conservative 81; Mismatches 192; Indels 227; Gaps 22;

QY	29	NLSGRWRNPDAGNVKAP-----GRDPLTCDEGDEPKSAHTSPKPKAKE-LPEMD	82
Db	610	NHMGWFCVEDD-GNHIKKEKHPLLVGHMPVMAKQOEKIECAAYNPFPYLNNESQPSDF	668
QY	83	ATS-SRWILSTLTTSKVIID-ANQPGTVVDQTCNAHVLCISSIPAAASDSYPFGE	140
Db	669	STAHGFLWIGSCTHQMGOIAIVSFQNSTPKVIECFNV-ESRILCMLYVPV-----	717
QY	141	MFLDSVNPDPGADGVLAGITLVGCATRCNVPNSCSRSGDTPVLDKGQGEVATIANGK	200
Db	718	-----EKKRREPGA-----	726
QY	201	VNPQSSTEEATEVDPGPEPETATLRPGPLTEHVFTDPATPSSGPQGSNGPEP	260
Db	727	-----PPDETFAVR-----	736
QY	261	DSSSTRPEPEPGDPTGAGSAAPTMMLGAQNGMYVHSVANWKKCLHSIKL-----	313
Db	737	-----ASDVPTICVTEGSIYKSSQGSKK-----VRLQHFPTPE	773
QY	314	KDSVLSLVHVKGRVLVALADGTALFHRGEGDQWDLNHYHMDLGHPPHHSIRCMVAVYDR	373
Db	774	KSTVMSLACTSQSLYAGLVNGAVASYARAPDGSMDSEPKVKILG--VLPVRSLLMMEDT	831
QY	374	VMCGYKKNKHVIOPTQWIEKSFDAHPRESQVQLAWIGDGVWVSRILDSLTIRLYHAHT	433
Db	832	LWAASGGQVFIISVETHAVEGLEAHQEBGVISHMAVSGVGIWIAFTSGSTLRLFTET	891
QY	434	HQHLQDVVDIEPYVKMLGTGKLGFSFVRITALLVAGSRWVGTGNGVVISIPLTETVLH	493
Db	892	LKHLQDINIATPVHMLP-----GHQRLSVTSLLVCHGLLVGTSLGVLVLPV-----	940
QY	494	RGQLLGRANKTSPTSSEGAPPGIIHVYGDSSDRAASSFIPYCSMAQAQLCFHGRDA	553
Db	941	-----PRLOGIPKVTG-----	963
QY	554	VKFPVSVPGNVLATL-----NGSVLDSPAEGPGPAAPASEVEGOKLRNVLSGEGYID	608
Db	964	VKFI-----VLATALHEKDKKSRDSLAPGPEP-----QDEDQK-----DALPSGGAG---	1006
QY	609	FRIGDGEDDET---EAGAGDMSQVKPVLSKAERS	639
Db	1007	SSLSQGDPAAIWLGDSIGSMTQ-KSDLSSSSGS	1039

RESULT 10

US-10-264-237-1721
 ; Sequence 1721, Application US/10264237
 ; Publication No. US20040009491A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Birse et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PA131P1
 ; CURRENT APPLICATION NUMBER: US/10/264,237
 ; CURRENT FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/16450
 ; PRIOR FILING DATE: 2001-05-18

```

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44323
; LENGTH: 52
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005920.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: SWISSPROT HIT: P34609, EVALUE 3.00e-09
; OTHER INFORMATION: EST_HUMAN HIT: BF131060.1, EVALUE 6.00e-26
; US-09-864-761-44323

Query Match 5.3%; Score 185.5; DB 9; Length 52;
Best Local Similarity 66.7%; Pred.No.3.1e-07;
Matches 36; Conservative 5; Mismatches 10; Indels 3; Gaps 2;

Qy 503 NKTSPSTGCGARGGGIHHVYGDSSDRAA-SSFIPYCSMAQAOLCFPHGRDAVK 555
Db 1 NKTSGVPGN--RPGSVIRVYGDNSDKVTPGTTFIPYCSMAHAQAOLCFPHGRDAVK 52

RESULT 12
US-10-114-270-98
; Sequence 98, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Zhaochong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Pedigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liete, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.

```

APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: No. US20040030110A1e1. Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 98
; LENGTH: 1809
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-98

Query Match
Best Local Similarity 4.6%; Score 161; DB 12; Length 1809;
Matches 138; Conservative 86; Mismatches 290; Indels 152; Gaps 28;

Qy 37 EDDAGNVKPA-----PGRDPLTCDREGDGPCKSAHTSPKCKAKELPEMDATSSR---- 87
Db 1228 DPERGSLAPAFSPSPAWIPVAPRAREAEKVPREERKSPKCKSMILSVLTSLQRPAGL 1287

Qy 88 --VMLTSTLTTSKVIIDANQPGTVVDQFTVCNAHVLCTISSIPAASDSDYPPGEMFLDS 145
Db 1288 IWHATSNQGFPSRLGGAEERPGT--DELAPMQSAVAE-PLPSPRAQPGG--TPA 1342

Qy 146 DVNP-----EDPGADVLGTLVGCATRCNVPRNSCSSRGDTPVLKGGQEVATIANG 199
Db 1343 DAGGQGSSEEP--ELVFA-----VNLPPAQLSSSD-----BETREELARI--G 1383

Qy 200 KVNPSQS--TEBATEATEVPDGPSEPETATLRPGPLTEHVFTDPAFTPSGPGQSGENGP 258
Db 1384 LVPPPEEFANGVLLATPLAGFGPS-PTTV---PSPASKPSEPPAPESAADSGVE--- 1436

Qy 259 EPDSSSTRPEPEPGDPTGAGSSAAPTWMLGAQNGWLY-VHSAVANWKKLIHSIKLKDSV 317
Db 1437 ---EADTRSSDPHLETTSTITVSSMSLTSSGELTDHTSFADG----HTFLEKPP 1489

Qy 318 LSLVHVHGRVLVALADGTIA----IFHRGEGQWDLNHYLMDLGHPHSHRCMAVYDR 373
Db 1490 ---VPPPKLKSPLGKGPVTRFDPLLKQSSSELMQAQHHASAG----- 1531

Qy 374 VWCYKKNKHVHQPKTQIEKSFDAHPRESQVRLAWIGDV-----WVSI 420
Db 1532 -----LASAAGFARPYLQFORSKLM-GDPVESRGLPGPEDDKPTVIS 1573

Qy 421 RLDSTLBYHANTHQHLOQVD-----IEPVYSXMLGTGKLGFSFVRITALLVAGSRILW 474
Db 1574 ELSRLAQLNKNDRSLGEEPVGGLGSLDDPAKKGPIAAARLFSSLGELSS--ISAQRSPG 1631

Qy 475 GTGNGVVISPLTETVTLHRCQLLGLRANKTSPISGARGPGGIIHVYDDSSDRAASSF 534
Db 1632 GPGGGASYVRPSGRYPVARRAPSPVPASLERVEGLGAGAGGAGRPFG-----L 1681

Qy 535 IPYCSMAQQLCFHGHRDAVKFFVSPGVNVLATLNGSLVLDSPAEGGPPAASEVEGQKL 594
Db 1682 TPPTTLKSSLSIPHEPKVRFVRSVSARSRSPPSPPLPSPASGPGPGAPGPRPFOOK 1741

Qy 595 RNVVLVSGGEGYIDFRIGD-----GE-----DDETEBAGDMSQVKPVLKASRSHII 642
Db 1742 PLQL-----WSKFDVGDWLESIHLEHRDFDEHEIEGAH-----LPALTKDDFVELG 1789

Qy 643 VMQVSY 648
Db 1790 VTRVGH 1795

RESULT 13
US-10-312-352-16
; Sequence 16, Application US/10312352
; Publication No. US20040053824A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
; APPLICANT: YUE, Henry; AZIMZAI, Yalda
; APPLICANT: HE, Ann; BATRA, Sajeev
; APPLICANT: LO, Terence P.; NGUYEN, Damiel B.
; APPLICANT: BURRILL, John D.; MARCUS, Gregory A.
; APPLICANT: ZINGLER, Kurt A.; GANDHI, Ameena R.
; APPLICANT: LAL, Preeti G.; KEARNEY, Liam
; APPLICANT: BURFORD, Neil; YAO, Monique G.
; APPLICANT: CHAWLA, Narinder K.; ELLIOT, Vicki S.
; APPLICANT: ARVIZU, Chandra S.; KHAN, Farrah A.
; APPLICANT: BAUGHN, Mariah R.; HAFALIA, April, J.A.
; APPLICANT: POLICKY, Jennifer L.; AU-YOUNG, Janice K.
; APPLICANT: LU, Yan; BOROWSKY, Mark L.
; APPLICANT: LU, Dyung Aina M.; RAMKUMAR, Jayalaxmi
; APPLICANT: YANG, Junming; GURURAJAN, Rajagopal
; APPLICANT: WARREN, Bridget A.; GIETZEN, Kimberly J.
; APPLICANT: XU, Yuming; KALLICK, Deborah A.
; APPLICANT: LEE, Ernestine A.; THANGAVELU, Kavitha
; APPLICANT: DELEGEANE, Angelo M.; LEE, Sally
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0794 USN
; CURRENT APPLICATION NUMBER: US/10/312,352
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: PCT/US01/21067
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,454
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 60/219,462
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 60/240,111
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: US 60/240,106
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: US 60/244,021
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/248,887
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/249,570
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 1829
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040053824A1 7477312CD1
US-10-312-352-16

Query Match 4.6%; Score 161; DB 12; Length 1829;
Best Local Similarity 20.7%; Pred. No. 0.0042;
Matches 138; Conservative 86; Mismatches 290; Indels 152; Gaps 28;

Qy 37 EDDAGNVKPA-----PGRDPLTCDREGDGPCKSAHTSPKCKAKELPEMDATSSR---- 87

Db 1248 DPERGSLASPAFSPRSPAWIPVAPRREAEKVRERKSPEDKSKMILSVLDTSLQRPAGL 1307
QY 88 --VMILSTLTTSKVVIIDANQPGTVVDQFTVCNAHVLCISSIPAASDSYPPGEMFLDS 145
Db 1308 IVVHATNGQFSLRGAEERPGT--PELAPAMQSAVAE-PLPSRAQPPGG--TPA 1362
QY 146 DVNP-----BDPGADVLAGITLVGCATRCNVPRNSCRRGDTFVLKDGQGEVATANG 199
Db 1363 DAGPQGSSEBEP--ELVFA-----VNLPPAQLSSSD-----BETREELARI--G 1403
QY 200 KVNFSQS--TEATEATEVDPGSPETATLRPGPLTEHVTDPAFTSSGPPQSGSENGP 258
Db 1404 LVPPPEEFANGVLATPLAGPGPS--PTTV-----PSPAGKPSSEPPAPESADSGVE--- 1456
QY 259 EPDSSTREPEPSGDPGTAGSAAPTMWLGAQNGWLY--VHSAVANWKKCLHSLIKDQSV 317
Db 1457 ---EADTRSSDPHLETTSTSTVSSMTLSSGELTDTHTSFADG---HTFLLEKPP 1509
QY 318 LSLVHVGRVLVALDGTIA-----IFHRGEGQWDLNHYHMLDGLPHHSIRCMVAVYDR 373
Db 1510 ---VPPKPKLKSPLGKGPVTRDPDLLKQSDSELMAQQHHAASAG----- 1551
QY 374 VWCYKKNVHVIOPTMQIEKSFDAHPRESQVRQLAWIGDGV-----WYSI 420
Db 1552 -----LASAAGPARPRYLFORRSKLW--GDVRESRGLPGPEDDXPTVIS 1593
QY 421 RLDSLRLYHAHTHOHQDQD-----TEPVYVKMLGTGKLGFSVRITALLVAGSRLMW 474
Db 1594 ELSRLQQLNKTDRSLGEEPPVGLGSLDDPAKKSPIAAARLFSSLGELSS--ISAQRSPG 1651
QY 475 GTGNGVLSIPLTETVVLHKGQLLGRANKTSPTSGEGARPGIITHVYGDSSDRAASGF 534
Db 1652 GPGGASYSVRPSGRYPVARRAPSPKASLVEREGLGAGAGACGRPPG-----L 1701
QY 535 LPYCSMAQAQLCFHGRDAVKEFVSVPGNVLATLNGSVLDSPAEGPGPAAPASEVEGOKL 594
Db 1702 TPPTILKSSLSIPIHEPKEVFRVVRVSARSRSPSPPLSPASGFGCAGPRPFQOK 1761
QY 595 RNVLVSGEGEYIDFRIGD-----GE-----DDETEBAGDMSQVKPVLKASRSHII 642
Db 1762 PLQL-----WSKFDVGDWLESIHLEHRDRFEDHEITGAH-----LPALTKDDFVELG 1809
QY 643 VWQVSY 648
Db 1810 VTRVGH 1815

RESULT 14
US-09-917-384-1
; Sequence 1, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,384
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1228
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
US-09-917-384-1

Query Match 4.4%; Score 154.5; DB 10; Length 1228;
Best Local Similarity 23.6%; Pred. No. 0.0078;
Matches 81; Conservative 27; Mismatches 112; Indels 123; Gaps 16;
QY 16 KDPTMKLWCAAG-----VNLSGWRPNEDDAGNGV-----KPAGRDPLTCD--REG 59
Db 288 KEPTMTATQVGGQPVESANFYQWNPDIDEADYAVDLYSRLVAAGFFSSIGMLIDTLRNG 347
QY 60 DGEPKSAHTSPEKKKAKELPEMDATSSRVWILSTLTTSKVVIIDANQPGTVVDQFTV-- 117
Db 348 WGGPNEP--TGP-----STAT-----DVN---TFVNSQSKIDL 374
QY 118 -----CNAHVLCISSIPAASDSYPPGEMFLDSVNPEDPG--ADGVLAG----- 160
Db 375 ROHRGLMCNQAGAGLQGPQASPTDFP--NAHLDAYVWIKPFGESDGTSAASDPTTGKKS 432
QY 161 -----ITLVGCAT-----RCNVRNSCNSRSGDTPVLDK 188
Db 433 DPMCDPTVTTSYGVLTNALPNSPIAGQWFPQAQDQLVANARPAVPTSTSSSPPPPP-- 488
QY 189 GQGEVATTANGKVPQSQTTEATEATEVDPGSPETATLRPGPLTEHVTDPAFTSS 248
Db 489 -----PSPSASPSPSPSPSPSPSPSPSPSPSPSPSPSPSPS--SSPSPSPS 537
QY 249 GPQSGSENGPSPDS-----STRPEPEPSGDPGTGAGS--SAAPT 285
Db 538 SPT 580
RESULT 15
US-09-917-383-1
; Sequence 1, Application US/09917383
; Publication No. US20030104522A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,383
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1228
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
US-09-917-383-1

Query Match 4.4%; Score 154.5; DB 10; Length 1228;
Best Local Similarity 23.6%; Pred. No. 0.0078;
Matches 81; Conservative 27; Mismatches 112; Indels 123; Gaps 16;
QY 16 KDPTMKLWCAAG-----VNLSGWRPNEDDAGNGV-----KPAGRDPLTCD--REG 59
Db 288 KEPTMTATQVGGQPVESANFYQWNPDIDEADYAVDLYSRLVAAGFFSSIGMLIDTLRNG 347
QY 60 DGEPKSAHTSPEKKKAKELPEMDATSSRVWILSTLTTSKVVIIDANQPGTVVDQFTV-- 117
Db 348 WGGPNEP--TGP-----STAT-----DVN---TFVNSQSKIDL 374
QY 118 -----CNAHVLCISSIPAASDSYPPGEMFLDSVNPEDPG--ADGVLAG----- 160
Db 375 ROHRGLMCNQAGAGLQGPQASPTDFP--NAHLDAYVWIKPFGESDGTSAASDPTTGKKS 432
QY 161 -----ITLVGCAT-----RCNVRNSCNSRSGDTPVLDK 188

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 11:07:28 ; Search time 44 Seconds
(without alignments)
1423.197 Million cell updates/sec

Title: US-10-019-495-9

Perfect score: 3473

Sequence: 1 MKNVFPVPCRLPEKEDPTM.....VLSKAERSHIIWQVSYTPE 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1152.5	33.2	1139	2 S40932	hypothetical prote
2	464.5	13.4	766	2 JCS958	sperm surface prot
3	306	8.8	988	2 T03307	hypothetical 109.6
4	142	4.1	300	2 J02220	hydroxyproline-ric
5	138.5	4.0	1032	2 T34433	hypothetical prote
6	138.5	4.0	2232	2 T34434	hypothetical prote
7	138	4.0	1013	2 C83771	hypothetical prote
8	138	4.0	1367	1 S48478	glucan 1,4-alpha-g
9	135	3.9	801	2 T29018	hypothetical prote
10	133	3.8	731	2 B86369	hypothetical prote
11	133	3.8	1469	2 T09219	hypothetical prote
12	132.5	3.8	1115	1 IQM5NL	basal transcriptio
13	132	3.8	1151	2 T18535	neural cell adhesi
14	131	3.8	677	2 J07303	high molecular mas
15	129	3.7	1183	2 T39233	pectate lyase (EC
16	129	3.7	1664	2 T18262	probable inositol
17	129	3.7	1779	2 T31085	S-layer protein -
18	128.5	3.7	585	2 E70930	xylanase - Caldice
19	128	3.7	311	2 T08781	hypothetical prote
20	128	3.7	929	2 C96623	hypothetical prote
21	127	3.7	473	2 S50755	hypothetical prote
22	126	3.6	880	2 D89756	hypothetical prote
23	125.5	3.6	839	2 F75518	protein T23E7.2b f
24	125.5	3.6	1506	2 T30886	hypothetical prote
25	125	3.6	309	2 T29293	integumentary muc
26	124.5	3.6	351	2 S50754	hypothetical prote
27	124.5	3.6	535	1 S76953	hypothetical prote
28	124	3.6	318	2 T35213	protein kinase (EC
29	124	3.6	330	2 T26004	probable secreted
					hypothetical prote

ALIGNMENTS

RESULT 1

S40932

hypothetical protein ZK1098.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997

C:Accession: S40932

R:Thomas, K.

submitted to the EMBL Data Library, February 1992

A:Reference number: S40923

A:Accession: S40932

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1139 <THO>

A:Cross-references: EMBL:Z22176; NID:g297978; PID:g297988

C:Genetics:

A:Introns: 29/3; 123/1; 183/2; 224/1; 280/2; 295/1; 307/1; 313/1; 397/3; 463/2; 501/2; 6;

Query Match

Best Local Similarity 33.2%; Score 1152.5; DB 2; Length 1139;

Matches 256; Conservative 110; Mismatches 202; Indels 99; Gaps 21;

Qy	1	MKNVFPVPCRLPEKEDPTMKLWCAAGVNLGWRPNEDDAGNVKPAEGRDPLTCDEGD	60
Db	542	VSSVPIPVCCRLPDNPEFLKWCATGVLRGGR---DERGQWIV---GDPIYFAPASM	594
Qy	61	GEPSAHTSPK-----KKAKEL-----DEMD--ATSSRWILTSTLTTSKVVIIDANQGT	110
Db	595	KKTKTSNHRPELEDEIKEARNLDAESELDEWQSSLVVWSSNGGKSLIAVLDAANNPN	654
Qy	111	VVDQFTVCNAHVLCTISSIPAAASDSYPPGE----MFLDSVDVNPED--PGADGVLAGITLV	164
Db	655	IIETFPACDHLCTIQAVSGVMEGEPMNERQSKYLGGGKI KDLPEGLDG----TDL	709
Qy	165	GCATRCNVPRNSCSRRGDTPLVDKQGEVATIANGKNVPSQSTEEATEATEVDPGPSEP	224
Db	710	GACSWELKRMWEDSDG-----VPTYCSNDMKPSKPTRDFSISEVAPVDSAP	758
Qy	225	ETATLRPGPLTEHVFTDPAPTPSGPGOPGSENGPEPDSSSTRPEPSGDDTGTGSSAAP	284
Db	759	-----VKEDPLPPAN--RPGGRAALPP---HIRDAMSKYDGVSGQMSGALP	800
Qy	285	TMWIGAQNGMLYHSAVANMKKCLHSIKLKDSVLSLVHVKGRVLVALADGTALIFHRGED	344
Db	801	TVMGGQNGQYIYIHSATANKQCLRRIKMPDVALSIVHYKSRIFAALANGTIALIFHRKH	860
Qy	345	GQWDLNHYHMLDLGPHHSIRCMVAVYDVRVWCVGKVKVHIQPKTMQIEKSFDAHPRES	404
Db	861	GENSDEGYHSLRVGSATSSVRSCLVSTNIWATYKNCVWVLDAESLQIVKVFAAHPRKDS	920
Qy	405	QVROLAMIGGVWVYSIRLDSITRLRYHAHTHOHLQDVIEPVVSKMLGTGKLGFSFVITA	464
Db	921	QVRNMQWVGAVLWSIRLDSITRLRYHAHTYEHQLQDVIEPVVTKMLGTSKLDFSYMRTTA	980

T8K14.10 (imported
probable membrane
extensin-like prot
protein CNK - frui
hypothetical 47.8K
hypothetical prote
hypothetical prote
probable iron-sulf
Doc4 protein, stre
hypothetical prote
hypothetical prote
procylicin acidic r
procylicin PSSA-1 -
CREB-binding prote
hypothetical prote
gastric mucin (Glo


```

Qy      285  TMMWIGA 290
      :
Db      185  APYYGA 190

RESULT 5
T34433
hypothetical protein K06A9.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C:Accession: T34433
R:Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid K06A9.
A:Reference number: Z21525
A:Accession: T34433
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1032 <GEI>
A:Cross-references: EMBL/U08046; PIDN:AACT0889.1; GSPDB:GN00028; CESP:K06A9.1a
A:Experimental source: strain Bristol N2; clone K06A9
C:Genetics:
A:Gene: CESP:K06A9.1a
A:Map position: X
A:Intons: 38/1: 75/3: 103/3: 132/2: 158/2: 222/1: 839/1: 848/1: 875/3: 890/2: 955/2:

```

```

Query Match          4.0%; Score 138.5; DB 2; Length 1032;
Best Local Similarity 25.86; Pred. No. 0.26;
Matches              67; Conservative 35; Mismatches 107; Indels 51; Gaps 13;

Qy      41  GNGVKPAPGDDPLTCDREGDGEPKSAHTSPEKKAKELPENDATSSRWILSTLT--S 98
       || : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db      320 GSTITPYFG---SSSTIGSGSTPSASSSSSGTMSTIS---GSTGTVTVVPGSSSFAS 371

Qy      99  KVIILDANQBCTVVDQFTVC-NAHLVICISIPAAUSDSDYPGCMFLDSVNPDPCGADGV 157
       || : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db      372 STPIASSSSPGSTV---TVAPGSSSTYGSSSTPSASSS--SGTMSTNS-----GSTGS 419

Qy     158  LAGITLVGCAT-RCNVPFSCNRGDTPVLVDKGGEVATIANGKNVFSQSTEATEATEV 216
       || : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db     420 TVTVAPVSSSTFGSGSTPIASSSSSGSTVTVVVG---SSSYGSSSTPSASSSAGTASTI 475

Qy     217  PDRGPSBPETALRPGPLETHVFDPAPTSP-----SGPQ-----PGSENGEPED 261
       ||| : | | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db     476 ---SGSTGTSTATIVPGSSSVSGSTQASPSPGCTMTVSGPTGSTVTVVPGSSTSPAPS 532

Qy     262  SSSTRPEPESGDPDTGAGSS 281

Db     533 SS-----PNPSSSPASTGST 547

```

RESULT 6
T34434
hypothetical protein K06A9.1a - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 29-Oct-1999 #sequence_revision 23-Oct-1999 #text_change 18-Feb-2000
C;Accession: T34434
R;Geisel, C.; Gattung, S.
submitted to the EMBL data Library, December 1996
A;Description: The sequence of *C. elegans* cosmid K06A9.
A;Reference number: Z21525
A;Accession: T34434
A;Status: preliminary; translated from GE/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2232 <GEI>
A;Cross-references: EMBL:U80846; PIDN:AA070890.1; GSPDB:GN00028; CESP:K06A9.1a
A;Experimental source: strain Bristol N2; clone K06A9
C;Genetics:
A;Gene: CESP:K06A9.1a
A;Map position: X
A;Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/1

```

Query Match      4.0%; Score 138.5; DB 2; Length 2232;
Best Local Similarity 25.8%; Pred. No. 0.72;
Matches 67; Conservative 35; Mismatches 107; Indels 51; Gaps 13

QY 41 GNGVKPAPGRDPLTCDRGDEGPEKSAHTSPKPKKAKELPEMDATSRVWILTSTLTT--S 98
Db 320 GSTITPVEG-----SSSTIGSSTPSASSSSSGTMSTIS-----GSTGSTVTVVPGSSSTTFAS 371
QY 99 KVIIDANQPGTVVDQFTVC--NAHVLCIHSSIPAAASDSYPPGEMFLDGDVNPEDPDGADV 157
Db 372 STPIASSSSPGSTV--TVAPGSSSTYGSGTSPASSSS--SGTMSTNS-----GSTGGS 419
QY 158 LAGITLVCCAT-RCNVPNSNCSSRGDTPVLKDGQGEVATIANKNVPQSGTEATEATEV 216
Db 420 TVTVAPVSSSTFGSGSTPIASSSSSGSTVTVVSG---SSSTYGSGTSPASSSSAGTASTI 475
QY 217 PDGPGSEPETATLRPGPLTEHVFTDPAPTPS-----SGPQ-----PGSENGPEPD 261
Db 476 ---SGSTGSTATIVPGSSSVGSGTQASPSPGTMTSVSGTGTSTVTVVPGSSTSPAPS 532
QY 262 SSSTRPEPEPGSDPTGAGSS 281
Db 533 SS-----PNPSSSPASTGST 547

RESULT 7
CB3771
hypothetical protein BH0971 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: CB3771
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii,
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodur-
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: CB3771
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1013 <STO>
A:Cross-references: GB:AF001510; GB:BA000004; NID:gi0173440; PIDN:BA04690.1; G
A:Experimental source: strain C-125
C:Genetics:
C:Gene: BH0971

```

Query Match	4.0%;	Score 138;	DB 2;	Length 1013;
Best Local Similarity	25.1%;	Pred. No. 0.28;		
Matches	79;	Conservative 36;	Mismatches 92;	Indels 108; Gaps 19;
Qy	138	PGEMFLDSVNPDPGA-DGVLAGI----	TLVGATRCNVPRSNCSRGDTPVLDKGQGV	193
Db	383	PGE-----	FORPGAPEDLMGLNPTGLDGMQEQINI-IK-----	EL 417
Qy	194	ATIANGVNPQSQTTEAT--EATEVPDPGP----	SEPTATLRPGPLTEHVFTDPDAPT	245
Db	418	LT---QLEPEK--EETPPEPTPPDPGPDPPEITEPVP----	PGP-----	DDPT 461
Qy	246	PSSGPQGSNGEPPEPOSSSTRPEPSGDPGTGAGSSAAP--	TMLGAQNGWLIVHSAVANW	304
Db	462	PDPEPTP-----EPEPDPT--PDPEPGEPEQEDDRSPYTIWERA-----	ALVEQL	509
Qy	305	KKCI-----HSIKLKDSVLIVHVKGRVLVALADGTIAI	FRGEDGQWDLSNY--	HLMDLG 358
Db	510	EKELGQKSNIKNNDL-----	VSNYDRILEEL--	537
Qy	359	HPHHSIRCMVAVYDRVWCYGKKNVHVIOPTMOIE--	KSFDAHPRESQVQRLAWIGDGV	416
Db	538	--LHYVDQLEKRDRDEVPNSLTIKIEREKKILSEPAEFASHFEPIQNRDPDKLMDYY		595
Qy	417	WVSRIDSTLRLYHA	431	
Db	596	GLLAELTMRITYHS	610	

RESULT 8

S48478
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - Yeast (*Saccharomyces cerevisiae*)
N/Alternate names: extracellular glucosylase; mucin-like protein MUC1; protein YIRO19c
C/Species: *Saccharomyces cerevisiae*
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 12-Nov-1999
C/Accession: S48478; A26877; B26877; S27281; JC6123
R/Rowley, K.
submitted to the EMBL Data Library, October 1994
A/Reference number: S48478
A/Accession: S48478
A/Molecule type: DNA
A/Residues: 1-1367 <ROW>
A/Cross-references: GB:Z47047; EMBL:Z38061; NID:g603364; GSPDB:GN00009; MIP
R/Yamashita, I.; Nakamura, M.; Fukui, S.
J. Bacteriol. 169, 2142-2149, 1987
A/Title: Gene fusion is a possible mechanism underlying the evolution of STA1.
A/Reference number: A91831; MUID:87194600; PMID:3106330
A/Accession: A26877
A/Molecule type: DNA
A/Residues: 1-242 <YAM>
A/Cross-references: EMBL:M16164; NID:g172522; PIDN:AAA35014.1; PID:g172525
A/Accession: B26877
A/Molecule type: DNA
A/Residues: 762-1331 <YA2>
A/Cross-references: EMBL:M16165; NID:g172523; PIDN:AAA35015.1; PID:g172526
R/Pardo, J.M.; Ianez, R.; Zalacain, M.; Claros, M.G.; Jimenez, A.
FEBS Lett. 239, 179-184, 1988
A/Title: Similar short elements in the 5' regions of the STA2 and SGA genes from *Saccharomyces cerevisiae*
A/Reference number: S27281; MUID:89031230; PMID:3141213
A/Accession: S27281
A/Molecule type: DNA
A/Residues: 1-31 <PAR>
A/Cross-references: EMBL:X13857; NID:g4551; PIDN:CAA32069.1; PID:g4552
R/Lambrechts, M.G.; Bauer, F.F.; Marmur, J.; Pretorius, I.S.
Proc. Natl. Acad. Sci. U.S.A. 93, 8419-8424, 1996
A/Title: Muc1, a mucin-like protein that is regulated by Mss10, is critical for pseudohyphae formation in *Saccharomyces cerevisiae*
A/Reference number: JG6123; MUID:96323237; PMID:8710886
A/Accession: JG6123
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-1367 <LAM>
A/Cross-references: GB:U30626; NID:g1304386; PIDN:AAC49609.1; PID:g1304387
C/Genetics:
A/Gene: SGD:MUC1; STA2; MAL5; DEX2; SGD:S0001458
A/Cross-references: MIP:S.YIRO19c; SGD:S0001458
A/Map position: 9R
C/Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase
C/Keywords: glycosidase; hydrolase; polysaccharide degradation; transmembrane protein
F:5-21/Domain: transmembrane #status predicted <TM1>
F:1350-1366/Domain: transmembrane #status predicted <TM2>

Query Match 4.0%; Score 138; DB 1; Length 1367;
Best Local Similarity 24.2%; Pred. No. 0.41;
Matches 60; Conservative 25; Mismatches 121; Indels 42; Gaps 8;

QY 47 ACRDPLTCRDGDGPKSAHTSPKKAKELPEMDATSSRVWILTSILTSKVLIIDAN 106
Db 456 APTVPTSSSTSSAPVTSSTSSAPVPTSSSTSSAPVTSSTSSAPVPTP 515
QY 107 QGTVDQFTVCNAHVLCITSSIPAAASDSYPPGEMFLDSVNPDPGADVLGTLVGC 166
Db 516 SSSITSS-----SSAPPT-----PSSSTSSAP-----VTSSTSS 550
QY 167 ATRCNVPRNCSRGDTPV-----LDKGQGEVATANGKVNPSQSTEEATEVDPGPGS 222
Db 551 SAPVPTFSSSTSSSTSTPTSSSTSSAPVPT-----PSSSTSSSA-PVPTPSSS 602
QY 223 EPEATL-RPGPLTEHVFTDPATPSSGPGQSGENGPEDSSSTR-----PEPEPGDPTG 277
Db 603 TTSSSAPAPTSSSTSSAPVTSSTSSAPVPTPSSSTSSAPVPTPSSSTTTE 662
QY 278 AGSSAAPT 285

Db

663 SSSAPVPT 670

: : : ||

RESULT 9

TS29018
hypothetical protein ZK84.1 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text_change 17-Mar-2000
C/Accession: TS29018
R/Kirsten, J.
submitted to the EMBL Data Library, April 1995
A/Description: The sequence of *C. elegans* cosmid ZK84.
A/Reference number: Z20553
A/Accession: TS29018
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-801 <KIR>
A/Cross-references: EMBL:U23181; PIDN:AAC48204.1; GSPDB:GN00020; CESP:ZK84.1
A/Experimental source: strain Bristol N2; clone ZK84
C/Genetics:
A/Gene: CESP:ZK84.1
A/Map position: 2
A/Introns: 22/2; 45/3; 108/1
C/Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match 3.9%; Score 135; DB 2; Length 801;
Best Local Similarity 22.6%; Pred. No. 0.32;
Matches 72; Conservative 37; Mismatches 143; Indels 66; Gaps 15;

QY 2 KNPVPVYCRPLVEKDPMTKLWCAAGVNLGWRPNEDDAGN--GVKPAFGRD-----PL 53
Db 282 EEAPAPA---PTAETPAPEVTSAAPEANSY----DSAGGDAATAPSSSAADAAPT 334
QY 54 TCDREGDGEKPSAHTSPKKAK--ELPEMDATSSRVWILTSILTSKVLIIDA-----NQ 107
Db 335 DSAASADTTAALVDTSSTSSHAESAESTPATDIAATE---TTPAPSPVPVADAAAGYDS 390
QY 108 PGTVDQFTVCNAHVLCITSSIPAAASDSYPPG---EMFLDSVNPDPGADVLGTL 163
Db 391 PSSIPPE-----TPAPAEEDTPAPASAAEETPAPAAEETPAPETASAPDA 439
QY 164 VGCAATRCNV---PRSNCSRGDTPVLDKGQGEVATANGKVNPSQSTEEAT-----EATRV 216
Db 440 AGGAAPADVAADVAADVAATTAPEITSSAQSAAGSYDV-----PSEPASEVTAIVESA 492
QY 217 PD-----PGPSEPETATLRPGPLTEHVFTDPATPSSGPGQSGENGPEDSSSTRPE- 268
Db 493 PSDSAAPITGAASEPAPAPIE-APATDAATLETPAPAAEPAPAAEAAAGYDAPSSVPEE 551
QY 269 ---PEPSGDPTGAGSSAA 283
Db 552 TPAPAPAADETTPAPAPAA 569

RESULT 10
B86369
hypothetical protein F508.10 - *Arabidopsis thaliana*
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence revision 02-Mar-2001 #text_change 31-Dec-2001
C/Accession: B86369
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A/Reference number: AB6141; MUID:21016719; PMID:11130712
A/Accession: B86369

A:Molecule type: mRNA
 A:Residues: 804-1081 <BA3>
 A:Cross-references: EMBL:X07244; NID:G53321; PIDN:CAA30230.1; PID:G929720
 R:Rougon, G.; Marshak, D.R.
 J. Biol. Chem. 261, 3396-3401, 1986
 A:Title: Structural and immunological characterization of the amino-terminal domain of m
 A:Reference number: A44290; MUID:86140120; PMID:3512556
 A:Accession: A44290
 A:Molecule type: protein
 A:Residues: 20-36 <ROU>
 C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol
 C:Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:IJMS
 C:Genetics:
 A:Gene: NCAM
 A:Map position: 9
 A:Introns: 643/3; 701/1; 770/2; 809/2; 1076/2
 C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu
 C:Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; si
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:120-1115/Product: neural cell adhesion molecule, long domain splice form #status experi
 F:720-809,1077-1115/Product: neural cell adhesion molecule, short domain splice form #sta
 F:720-711/Domain: extracellular #status predicted <EXT>
 F:734-98/Domain: immunoglobulin homology <IMM1>
 F:132-191/Domain: immunoglobulin homology <IMM2>
 F:152-156/Region: heparin binding #status predicted
 F:161-165/Region: heparin binding #status predicted
 F:228-290/Domain: immunoglobulin homology <IMM3>
 F:262-272/Region: NCAM binding #status predicted
 F:323-388/Domain: immunoglobulin homology <IMM4>
 F:420-482/Domain: immunoglobulin homology <IMM5>
 F:519-596/Domain: fibronectin type III repeat homology <FN3A>
 F:623-625/Domain: fibronectin type III repeat homology <FN3B>
 F:712-729/Domain: transmembrane #status predicted <TM>
 F:730-1115/Domain: intracellular #status predicted <INT>
 F:41-96,139-189,235-288,330-386,427-480/Disulfide bonds: #status predicted
 F:222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 3.8%; Score 132.5; DB 1; Length 1115;
 Best Local Similarity 22.5%; Pred. No. 0.73;
 Matches 83; Conservative 40; Mismatches 127; Indels 119; Gaps 17;
 QY 22 LWCAAGVNLGWRPNEDDAGNGVKPAPGRD-----PL-----TCDEGDGE 62
 DB 742 LMCIA-VNLG-----KAGPAK--GKMEGKAAFSKDEKEPIVEVRTEERTPNH 791
 QY 63 PKSAHTSPEKKAKELPEM--DATSRVWLTS--TLTTSKVIIIDA-----NQP----- 108
 DB 792 DGGKHEPNETTPELPADTTATVEDMLPSVTVTNDSITETATQNSPTSETT 851
 QY 109 -----GTVVQFTV-----CNAHVLICISIPAAASDSYPPGEMFLDSVNPEDP 152
 DB 852 TLTTSIAPPATTVPDSNSVPAGATFSKGVTTASSSPASAPKVAELVDL---SDPTTSAP 908
 QY 153 GAD-----GVLAGITLVGCATRCNVPRNSCSRGDTPVLDKGGQEVATIANGKV 201
 DB 909 SASNLSSITVLANGQVLSPTASAGETSKAPPASKASPATPTPAGAASPLAAVAAPAT 968
 QY 202 NPSQSTEATEATEVDDPQSE-----PETATLPGLTEHVFTDP----- 242
 DB 969 DAPQAKQEA-PSTKGPDPPEFTQGTVPKNPEAATAPSPKSAATNTPSQGEDLKMDEGN 1027
 QY 243 -----AFTSSGPO-PGSENGPEPDSSTTRPEP-----EFSGDPT 276
 DB 1028 FKTPDIDLAKOVFAALGSRPAPGASQASQASAPADSAVPPAKTEKGVETKSEPP 1087
 QY 277 GAGSSAAPT 285
 DB 1088 ESEAKPAPT 1096
 RESULT 13
 T18535
 high molecular mass nuclear antigen - chicken (fragment)

C:Species: Gallus gallus (chicken)
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T18535
 R:Shimada, K.; Harata, M.; Mizuno, S.
 J. Cell Sci. 110, 3031-3041, 1997
 A:Title: A nuclear matrix-associated high molecular mass nuclear antigen, HMNA, of chicken
 A:Reference number: Z18955; MUID:9803440; PMID:9365273
 A:Accession: T18535
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1151 <SHI>
 A:Cross-references: EMBL:D89440; NID:d1177136; PID:d1025045; PIDN:BAA24137.1
 Query Match 3.8%; Score 132; DB 2; Length 1151;
 Best Local Similarity 22.5%; Pred. No. 0.83;
 Matches 73; Conservative 29; Mismatches 129; Indels 94; Gaps 12;
 QY 45 KPAPGRDPLTCDREGDGPKSAHTSPEKKAKELPEMDATSRVWLTSILT----- 96
 DB 247 KPAPVTSPTIFCSSAAEKPLTAASPTASKATAEAKVPATAS---LMATKVTAEAKPAPS 303
 QY 97 -----TSKVIIIDANQPGTVVDQFTVCNAHVLCISS-----IPAAASDSYPGGE 140
 DB 304 PSVPKATTTDKAVTATAPKAGPD---VKPAVAVCAEAKPAPPPPPQQLPKAAAAAPTGT 360
 QY 141 MELSDVNPE-DRGADGVLGIT--LVGCATRCNVPRNSC---SSRGDTPVLDKQG--- 190
 DB 361 ELKPATAPPHSGSRANSHVTVTTPNVPRAAAAVPTAGAVPKASTGTTTAAAPQOPVPK 420
 QY 191 -----GEVATIANGKVNPSQSTEEA---TEATEVPDPGPSEPETATLRP--- 231
 DB 421 AAPVTPSPQOAVPRAATAAAAPVTPQPVTKAATTNATPPQPIPKAATTTTATPTVP 480
 QY 232 -----GPLETHVFTD---PAPTSSGPGQSGENGPEPD 261
 DB 481 QQPIPKAGTDAAPPAPVAPKAPSGDGAATPGVFNATDPKPPPTQSPVSAVTEPKQP- 539
 QY 262 SSSRTRPEPEPSGDDPTGAGSSAAPT 286
 DB 540 ----RAAPPNSNEATPAVPSFSPNL 560

RESULT 14
 JC7303
 pectate lyase (EC 4.2.2.2) - Bacillus sp.
 N:Alternate names: pectate transeliminase
 C:Species: Bacillus sp.
 C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
 C:Accession: JC7303
 R:Ogawa, A.; Sawada, K.; Saito, K.; Hakamada, Y.; Sumitomo, N.; Hatada, Y.; Kobayashi, T.
 Biosci. Biotechnol. Biochem. 64, 1133-1141, 2000
 A:Title: A new high-alkaline and high-molecular-weight pectate lyase =66 from a Bacillus
 A:Reference number: JC7303
 A:Molecule type: DNA
 A:Residues: 1-677 <OGA>
 A:Cross-references: DDBJ:AB028878
 A:Experimental source: strain KSM-PL5
 C:Genetics:
 A:Gene: pel-15H
 C:Keywords: calcium binding; carbon-oxygen lyase
 Query Match 3.8%; Score 131; DB 2; Length 677;
 Best Local Similarity 18.3%; Pred. No. 0.48;
 Matches 108; Conservative 86; Mismatches 214; Indels 182; Gaps 26;
 QY 45 KPAPGRDPLTCDREGDGPKSAHTSPEKKAKELPEMDATSRVWLTSILTTSKVIIID 104
 DB 95 QPLTGMVTVPQFDWYHEGNARASSLRPMK-----IFSASSESSSTTIVEIQ 139
 QY 105 ANQPGTVVDQFTVCNAHVLCISSIPAAASDSYPPGEMFLDSVNPEDPGADGVLGITLV 164
 DB 140 TRDGGRHIAQ-TVNGTHLLV-----TDFAPDTWY----- 168


```
QY 165 GCATRCNVPRNSCSSRGTDPVLDKQGEVATIANGKVNPSQSTEBATEATEVDPGPSEP 224
Db 169 -----RFINVMTDTKKVD-----TYNGELKLEQANFASTSAQVORLKIYQ 212
QY 225 ETATLRPGPLTEHVF-----TDPATPSSG-POPGSE--NGPEPDSSTRP----- 267
Db 213 NSPTI--GOYIDNLYVYSGSNPCTDPGPGSPGTPDGTGTPGDPDPGSPDGNPNP 270
QY 268 ----BEPGSDPTGAGSSAPTMWLGQNGWLVHSAVANWKKLHSIKLDSVLSLVHV 323
Db 271 GHPQDPGPAPEAEGDLIVAP-----NGQEG-----NPGTLNQPTLTS-----AITRI 314
QY 324 KGRVLVALADGTIA-----IFHRGEG-----OMDLNHYHLM 355
Db 315 QPGRTIYMRGGTYAFSETVLIERNGNLEGARKIRVGVNGEKPVLDFSAQAFDPNRLQ 374
QY 356 DLGHPHHSIRCMVV-----YDRVCGVKN-KVHIQPKTMOIEKSFDAHPR 402
Db 375 INGHYWH-VQGIETVKEAGDNGIFIGGNYNRI-----ENVETHNKDTGLQISRYSSATRD 429
QY 403 ESQVROLAWIGDGVWVSIRLDSTLRLYHAHTHQLQDVDIIEPVVSKML-GTKLGFSEVR 461
Db 430 E-----WFS--YNEIIVYSHNNYDDPDGEDADGFAAKLTSGPGNV-FDGC 473
QY 462 ITALLVAGSRLWVGTVNGVVISIPLTETVHLRGQLGLRANKTSPTSCEGARPGIIVH 521
Db 474 AAYNVDDGMDLYTKSDTGAIYVIRNSIAYNNGSTEG--GHSTNSDNGGFKLG----- 527
QY 522 YGDDSSDRAASFIPYCSMAQQLCFHGHDRDAVKFFVSPGVNVLATINGS 571
Db 528 -----SNIPVNHIVENMAF-GNKKHGFTYNSPGSITMTNNTS 565

RESULT 15
T39233
Probable Inositol polyphosphate phosphatase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T39233
R:Churcher, C.M.; Gentles, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21837
A:Accession: T39233
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1183 <CHU>
A:Cross-references: EMBL:Z98763; PIDN:CAB11494.1; GSPDB:GN00066; SPDB:SPAC9G1.10c
A:Experimental source: strain 972h-; cosmid c9G1
C:Genetics:
A:Gene: SPDB:SPAC9G1.10c
A:Map position: 1

Query Match 3.7%; Score 129; DB 2; Length 1183;
Best Local Similarity 19.6%; Pred. No. 1.4;
Matches 118; Conservative 81; Mismatches 272; Indels 130; Gaps 25;

QY 7 PVYCRPLVEKDPWKMLWCAAGVNLGWRPNEDDAGNVKVPAPGRDPLTCDEGDGEPKSA 66
Db 138 PLLKRP-QQKGPESIFQSSVOSTGNDLMKHDDTNHQPPI--KPNFSSKAGSSSPIS- 193
QY 67 HTSPEKKAKELPMDATSSRWLIITLTTSKVVIIDANQPG-TVVDQFTVCNAHVLCI 125
Db 194 -VSPKLVNKAIVISQPTHEASSVLSSEEEENINSKVSFSLDHPDPSQTFGKECPI 252
QY 126 SSIFAA-----SDSDYPPGEMFLSDVNPEDPGADGVLAGITLVGCATRCNVP----- 173
Db 253 STAPPVLNIGRSLRLETPPP---IPSPRPQPVAVEAICQSRAVISQQLPLHVSPRKPKP 309
QY 174 --RNSCSRGDTPVLDKQGEVATIANG-KVNPSQSTEBATEATEVDP-----PGSEPE 225
Db 310 PLRKVSTQRSSSPIENLATKSDASLVLTGLQSSPYTHIAPASEMSLIPEKPLRPSPHTL 369
```

```
QY 226 TATLRPGPLTEHVFDTDPAPT-PSSGPOGSENGPEPDSSTRPEPEPSPGDPPTGAGSSAAP 284
Db 370 SELSGPALTSENLSSKPSPLFPFPPPRVKSLATNKEVS--MPVSTEQSDPSVAASSSSS 426
QY 285 TMWLGAQNGWLVHSAVANWKKC-----LHSIKLDSVLSLVHVKGRVLVALADGTIAIFH 340
Db 427 SQLDVLVKGSIPTDTSVRNPPCFVNGVESINV-DFEARI FVSGDRLLVLAGGGLRVY- 484
QY 341 RGEDQWDLNHYHLMDLGHPHHSIRCMVYVDRVMCGYKXNVHVIQPKTMOIEKSFDAHP 400
Db 485 ----DTVTGLCHWH-MPLG-----DIKVTLSLFSKSSP 511
QY 401 RRESQVROLAWIG--DG-VWVSIRLDSTLRLYHAHTHQLQDVDIIE--PYVSKMLGTGK 454
Db 512 ENYSDDGRFVWFGRDGMILW-----EVDVQNHIIVTKKSXSVSNCPITYVMVYKXNEMWTLDD 566
QY 455 LGFSFV-----RITALLVAGSR-LWVGTVNGVVISIPLTETV 491
Db 567 MGKLYVWQEDETMGLSIOSTPHSIRTIPTAHAMVLDNELLWVVVGKSIYVVDPTS--- 623
QY 492 LHRGQLGLRANKTSPTSCEGARPGIIVHVGDDSSDRAASFIPYCSMAQQLCFHGH 551
Db 624 -----ENESASVLAKDMPPLI---GDISCGTTISNF-----TDLVIFYGHV 662
QY 552 D 552
Db 663 D 663
```

Search completed: August 23, 2004, 11:15:09
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 11:04:17 ; Search time 25 seconds
(without alignments)
1355.907 Million cell updates/sec

Title: US-10-019-495-9

Perfect score: 3473

Sequence: 1 MKNVPVPVYCRPLVEKDPTM.....VLSKAERSHIIWQVSYTPE 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3473	100.0	1334	1	JIP3_HUMAN
2	3191	91.9	1337	1	JIP3_MOUSE
3	1217.5	35.1	1227	1	JIP3_DROME
4	1152.5	33.2	1157	1	JIP CAEL
5	183.5	5.3	1222	1	YNG2 CAEL
6	161	4.6	797	1	SHK3_HUMAN
7	146.5	4.2	1815	1	SHK3_RAT
8	138	4.0	1367	1	AMYH_YEAST
9	137.5	4.0	634	1	HWPI_CANAL
10	136	3.9	3644	1	MINT_MOUSE
11	132.5	3.8	1115	1	NCA1_MOUSE
12	130	3.7	677	1	OGFR_HUMAN
13	129	3.7	1664	1	SLP1_CLOTH
14	128	3.7	1567	1	FNW2_MOUSE
15	126.5	3.6	742	1	PKWA_THRCU
16	125.5	3.6	442	1	SDC3_MOUSE
17	124.5	3.6	535	1	SPKC_SYNY3
18	123.5	3.6	824	1	JIP2_HUMAN
19	123.5	3.6	886	1	SM6B_MOUSE
20	122.5	3.5	279	1	Y091_NPVOF
21	122	3.5	1225	1	DAT1_HUMAN
22	121.5	3.5	143	1	PARC_TRYBB
23	121.5	3.5	145	1	PARC_TRYBB
24	121.5	3.5	2441	1	CBP_MOUSE
25	121.5	3.5	5147	1	PCLO_HUMAN
26	120.5	3.5	565	1	MOT8_MOUSE
27	120.5	3.5	1134	1	ANS1_HUMAN
28	119.5	3.4	667	1	SIX5_MOUSE
29	119.5	3.4	1520	1	ACFD_ECOLI
30	119	3.4	720	1	2408_HUMAN
31	119	3.4	2415	1	PCCA_HUMAN
32	118.5	3.4	797	1	VGLX_HSVBB
33	118.5	3.4	1101	1	GUNC_CELFI

34 118 3.4 977 1 BAB1_DROME
35 117.5 3.4 1217 1 AFA_MOUSE
36 117.5 3.4 1742 1 GUN4_CALSA
37 117 3.4 1394 1 CNG4_BOVIN
38 116.5 3.4 385 1 YL90_MYCTU
39 116.5 3.4 1229 1 P121_HUMAN
40 116.5 3.4 1395 1 CUT1_MOUSE
41 116 3.3 387 1 T122_MOUSE
42 116 3.3 1276 1 SCAP_CRIGR
43 115.5 3.3 555 1 GPI_CHLRE
44 115.5 3.3 569 1 TACT_HUMAN
45 115.5 3.3 730 1 MM09_MOUSE

ALIGNMENTS

RESULT 1
JIP3_HUMAN
ID JIP3_HUMAN STANDARD; PRT; 1334 AA.
AC Q9UPT6; Q9ERY4; Q9H414; Q9H7P1; Q9NUG0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE C-Jun-amino-terminal kinase interacting protein 3 (JNK-interacting protein 3) (JIP-3) (JNK MAP kinase scaffold protein 3) (Mitogen-activated protein kinase 8-interacting protein 3).
GN MAPK8IP3 OR JIP3 OR KIAA1066.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=21096910; PubMed=11157797;
RX Danielis R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K., Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J., Higgs D.R.;
RT "Sequence, structure and pathology of the fully annotated terminal 2 Mb of the short arm of human chromosome 16.";
RL Hum. Mol. Genet. 10:339-352(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX Bagguley C., Hall R.;
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A., AND REVISIONS TO 322; 412 AND 751.
RX TISSUE=Brain;
RT Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 266-1334 FROM N.A.
RX TISSUE=Brain;
RT MEDLINE=9397452; PubMed=10470851;
RX Kikuno R., Nagase T., Ishikawa K.-I., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
RL DNA Res. 6:197-205(1999).
RN [5]
RP SEQUENCE OF 1-252 FROM N.A.
RX TISSUE=Spleen;
RT "The nucleotide sequence of a long cDNA clone isolated from human spleen.";
RN Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: The JNK-interacting protein (JIP) group of scaffold proteins selectively mediates JNK signaling by aggregating specific components of the MAPK cascade to form a functional JNK signaling module. May function as a regulator of vesicle transport, through interactions with the JNK-signaling components and motor proteins (By similarity).

Q9w0k7 drosophila
O88573 mus musculus
P22534 caldocellum
Q28181 bos taurus
Q10383 mycobacteri
Q9Y2n3 homo sapien
P53564 mus musculus
Q9eqn3 mus musculus
P97260 cricetus
Q9fpq6 chlamydomon
P40200 homo sapien
P41245 mus musculus

RA Kelkar N., Gupta S., Dickens M., Davis R.J.;
 RT "Interaction of a mitogen-activated protein kinase signaling module
 with the neuronal protein JIP3";
 RL Mol. Cell. Biol. 20:1030-1043(2000).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1C), SUBCELLULAR LOCATION, AND INTERACTION
 WITH KLC1.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=20560743; PubMed=11106729;
 RA Bowman A.B., Kamal A., Ritchings B.W., Philp A.V., McGrail M.,
 RA Gindhart J.G., Goldstein L.S.B.;
 RT "Kinesin-dependent axonal transport is mediated by the Sunday Driver
 (SYD) protein.";
 RL Cell 103:583-594(2000).
 RN [5]
 RP SEQUENCE OF 1240-1337 FROM N.A.
 RX MEDLINE=22388257; PubMed=1247932;
 RA Strausberg K.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP INTERACTION WITH KLC.
 TISSUE=Brain;
 RX MEDLINE=21135887; PubMed=11238452;
 RA Verhey K.J., Meyer D., Deehan R., Blenis J., Schnapp B.J.,
 RA Rapoport T.A., Margolis B.;
 RT "Cargo of kinesin identified as JIP scaffolding proteins and
 associated signaling molecules.";
 RL J. Cell Biol. 152:959-970(2001).
 CC -!- FUNCTION: The JNK-interacting protein (JIP) group of scaffold
 CC proteins selectively mediates JNK signaling by aggregating
 CC specific components of the MAPK cascade to form a functional JNK
 CC signaling module. May function as a regulator of vesicle
 CC transport, through interactions with the JNK-signaling components
 CC and motor proteins.
 CC -!- SUBUNIT: Forms homo- or heterooligomeric complexes. The central
 CC region of Mapk8ip3 interacts with the C-terminal of Mapk8ip2 but
 CC not Mapk8ip1. Binds specific components of the JNK signaling
 CC pathway namely Mapk8, Mapk9 and Mapk10 to the N-terminal region,
 CC Map2k4 and Map2k7 to the central region and Mapk11 to the C-
 CC terminal region. Binds the TPR motif-containing C-terminal of
 CC kinesin light chain, pre-assembled Mapk8ip1 scaffolding complexes
 CC are then transported as a cargo of kinesin, to the required
 CC subcellular location.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; localised in the soma and
 CC growth cones of differentiated neurites and the Golgi and vesicles
 CC of the early secretory compartment of epithelial cells.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=6;
 CC Name=1c; Synonyms=3b;
 CC IsoId=Q9ESN9-1; Sequence=Displayed;
 CC Name=1a;
 CC IsoId=Q9ESN9-2; Sequence=VSP_002775, VSP_002777;
 CC Name=1b;
 CC IsoId=Q9ESN9-3; Sequence=VSP_002776, VSP_002777;
 CC Name=1d;
 CC IsoId=Q9ESN9-4; Sequence=VSP_002778, VSP_002779;
 CC Name=3a;
 CC IsoId=Q9ESN9-5; Sequence=VSP_002778, VSP_002779;
 CC Name=1e;
 CC IsoId=Q9ESN9-6; Sequence=VSP_002776;
 CC -!- TISSUE SPECIFICITY: Highly expressed throughout many regions of
 CC the brain and at lower levels in the heart, liver, lung, testes
 CC and kidney. All isoforms have been identified in the brain,
 CC Mapk8ip3A is also expressed in the spleen and lung.
 CC -!- INDUCTION: Expressed in neurites 5 days following initiation of
 CC nerve growth factor Ngf induced differentiation. Ngf withdrawal
 CC results in the down-regulation of Mapk8ip3 protein by caspase-
 CC mediated cleavage.
 CC -!- SIMILARITY: Belongs to the JIP scaffold family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB005662; BAA5874.1; -
 CC EMBL; AB043124; BAB16675.1; -
 CC EMBL; AB043125; BAB16676.1; -
 CC EMBL; AB043123; BAB16674.1; -
 CC EMBL; AB043129; BAB16685.1; -
 CC EMBL; AB043126; BAB16685.1; JOINED.
 CC EMBL; AB043127; BAB16685.1; JOINED.
 CC EMBL; AB043128; BAB16685.1; JOINED.
 CC EMBL; AF178637; AAF26843.1; -
 CC EMBL; AF178636; AAF26842.1; -
 CC EMBL; AF262046; AAG36931.1; ALT_INIT.
 CC EMBL; BC004003; AAH04003.1; -
 CC MGD; MGI:1353598; Mapk8ip3.
 CC GO; GO:0005737; C:cytoplasm; IEP.
 CC GO; GO:0019894; F:kinesin binding; IPI.
 CC GO; GO:0005078; F:MAP-kinase scaffold activity; IPI.
 CC GO; GO:0019901; F:protein kinase binding; IPI.
 CC GO; GO:0046328; P:regulation of JNK cascade; IDA.
 CC GO; GO:0016192; P:vesicle-mediated transport; IDA.
 CC Alternative splicing; Phosphorylation; Coiled coil.
 CC DOMAIN 58 177
 CC COILED COIL (POTENTIAL).
 CC COILED COIL (POTENTIAL).
 CC MOD RES 266 266
 CC PHOSPHORYLATION (BY MAPK).
 CC MOD RES 276 276
 CC PHOSPHORYLATION (BY MAPK).
 CC MOD RES 287 287
 CC PHOSPHORYLATION (BY MAPK).
 CC VARSPLIC 201 201
 CC Missing (in isoform 1a and isoform 1d).
 CC /FTid=VSP_002775.
 CC VARSPLIC 201 201
 CC S -> SPRQSWRKS (in isoform 1b and isoform
 CC 1e).
 CC /FTid=VSP_002776.
 CC Missing (in isoform 1a and isoform 1b).
 CC /FTid=VSP_002777.
 CC Missing (in isoform 3a).
 CC /FTid=VSP_002778.
 CC Missing (in isoform 3a).
 CC /FTid=VSP_002779.
 CC R-SG: RESULTS IN INHIBITION OF JNK
 CC BINDING.
 CC P-SG: RESULTS IN INHIBITION OF JNK
 CC BINDING.
 CC T-SG: RESULTS IN INHIBITION OF JNK
 CC BINDING.
 CC S-SG: RESULTS IN INHIBITION OF JNK
 CC BINDING.
 CC L-SG: RESULTS IN INHIBITION OF JNK
 CC BINDING.
 CC T-A: RESULTS IN LOSS OF PHOSPHORYLATION
 CC OF MAPK8IP3; WHEN ASSOCIATED WITH A-276
 CC AND 287. DOES NOT EFFECT BINDING OF
 CC COMPONENTS OF THE JNK PATHWAY.

FT	MUTAGEN	276	276	T->A: RESULTS IN LOSS OF PHOSPHORYLATION OF MAPKIP3; WHEN ASSOCIATED WITH A-266 AND A-287. DOES NOT EFFECT BINDING OF COMPONENTS OF THE JNK PATHWAY.
FT	MUTAGEN	287	287	T->A: RESULTS IN LOSS OF PHOSPHORYLATION OF MAPKIP3; WHEN ASSOCIATED WITH A-266 AND A-287. DOES NOT EFFECT BINDING OF COMPONENTS OF THE JNK PATHWAY.
FT	CONFLICT	312	312	K -> R (IN REF. 4).
FT	CONFLICT	376	376	K -> L (IN REF. 3; AAF26843).
FT	CONFLICT	561	561	E -> K (IN REF. 4).
Query Match 91.9%; Score 3191; DB 1; Length 1337;				
Best Local Similarity 91.2%; Pred. No. 2.9e-187;				
Matches 594; Conservative 24; Mismatches 33; Indels 0; Gaps 0;				
Qy	1	MKNVPVPCPLVEKDPMTKLWCAAGVNLGWRNEDDAGNGVPAKPRDPLTCDREGD	60	
Db	687	MKNVPVPCPLVEKDPMTKLWCAAGVNLGWRNEDDAGNGVPAKPRDPLTCDREGD	746	
Qy	61	GEPSAHTSPKKAKELPEMDATSSRWILTTSTLTTSKVLIIDANQPGTVVDQFTVCNA	120	
Db	747	GEPSKTHSPKKAKETPEADATSSRWILTTSTLTTSKVLIIDANQPGTIVDQFTVCNA	806	
Qy	121	HVLCISSTPAASDSNPCEMFLSDVNPDPGADVLGTLVGCATRCNVPNSCNSR	180	
Db	807	HVLCISSTPAASDSNPCEMFLSDVNPDPGADVLGTLVGCATRCNVPNSCNSR	866	
Qy	181	GDTPLVDKGGEVATANGKNVPSOSTEATEATVPDPGPSEPETALRPGPLTEHVFT	240	
Db	867	GDTPLVDKGGEVATANGKNVPSOSTEATEATVPDPGPSESEATVTRPGPLTEHVFT	926	
Qy	241	DPAPTPSGPGSGNSGPEPSSSTRPEPESGDPGTAGSSAAPFMWLGQNGWLYVHSA	300	
Db	927	DPAPTPSSSTQPAESNGESNGTIVQPQVPSGELSTTTSSAAPFMWLGQNGWLYVHSA	986	
Qy	301	VANWKKCLHSIKLSDVLSLVHVKGRVLVALADGTLAIFHRGEDGQWDLNVLHMDLGH	360	
Db	987	VANWKKCLHSIKLSDVLSLVHVKGRVLVALADGTLAIFHRGEDGQWDLNVLHMDLGH	1046	
Qy	361	HHSIRCMVAVYDRVWCGYKKNVHVITQPKMTQTEKSFDAHPRESOVRLAWIGDGWVYSI	420	
Db	1047	HHSIRCMVAVNDRVWCGYKKNVHVITQPKMTQTEKSFDAHPRESOVRLAWIGDGWVYSI	1106	
Qy	421	RLDSTRLYHANTHOHLQDVIEPVVSKMLGTGKLGFSFVRITALLVAGSLRWGTGNGV	480	
Db	1107	RLDSTRLYHANTHOHLQDVIEPVVSKMLGTGKLGFSFVRITALLVAGSLRWGTGNGV	1166	
Qy	481	VISIPLTETVLIHRGQLGLRANKTSPTSGEGRPGIIVHVGDDSDRAASSFIPYCSM	540	
Db	1167	VISIPLTETVLIHRGQLGLRANKTSPTSGEGRPGIIVHVGDDSDRAASSFIPYCSM	1226	
Qy	541	AAQQLCFHGRDAVFFVSGNVLATLNGSVLDSPAEGPGPAAPASEVEGQKLRNLVL	600	
Db	1227	AAQQLCFHGRDAVFFVSGNVLATLNGSVLDSPAEGPGPAAPADAEGQKLRNLVL	1286	
Qy	601	SGGEGYIDPRIDGDEDETEGAGDMSQVPLSKAERSHIIVQVSYTPE	651	
Db	1287	SGGEGYIDPRIDGDEDETEGAGDMSQVPLSKAERSHIIVQVSYTPE	1337	
RESULT 3				
JIP3 DROME				
ID	JIP3 DROME	STANDARD;	PRT;	1227 AA.
AC	Q9GQF1; Q9VSK0; Q9VSC0;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	JNK-interacting protein 3 (Sunday driver protein).			
GN	SYD OR CG8110.			
OS	Drosophila melanogaster (fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			

OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A. (ISOFORM LONG), AND INTERACTION WITH KLC.
RX	MEDLINE=20560743; PubMed=11106729;
RA	Bowman A.R., Kamai A., Ritchings B.W., Philip A.V., McGrail M.,
RA	Gindhart J.G., Goldstein L.S.B.;
RA	"Kinesin-dependent axonal transport is mediated by the sunday driver
RT	(SYD) protein.";
RL	Cell 103:583-594 (2000).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Berkley;
RX	MEDLINE=20196006; PubMed=10731132;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
RA	Abriell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA	Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA	Fosler K.G., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hoskins D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA	Jakobi M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mouton S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Neus D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT	"The genome sequence of Drosophila melanogaster.";
RL	Science 287:2185-2195 (2000).
RN	[3]
RP	REVISIONS, AND ALTERNATIVE SPLICING.
RX	MEDLINE=22426069; PubMed=12537572;
RA	Misra S., Crosby M.A., Murgall C.J., Matthews B.B., Campbell K.S.,
RA	Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA	Smith C.D., Tupy J.L., Whitfield B.J., Bayraktaroglu L., Berman B.P.,
RA	Bettencourt B.P., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA	Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA	Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA	Lewis S.E.;
RT	"Annotation of the Drosophila melanogaster euchromatic genome: a
RT	systematic review.";
RL	Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN	[4]
RP	SEQUENCE OF 937-1227 FROM N.A.
RC	STRAIN=Berkley; TISSUE=Head;
RX	MEDLINE=22426066; PubMed=12537569;
RA	Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA	George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA	Rubin G.M., Celniker S.E.;
RT	"A Drosophila full-length cDNA resource.";

Genome Biol. 3:RESEARCH0080.8(2002).
 CC !- FUNCTION: The JNK-interacting protein (JIP) group of scaffold
 CC proteins selectively mediates JNK-signaling by aggregating
 CC specific components of the MAPK cascade to form a functional JNK
 CC signaling module. May function as a regulator of vesicle
 CC transport, through interactions with the JNK-signaling components
 CC and motor proteins. Syd is required for efficient kinesin-I
 CC mediated axonal transport.
 CC !- SUBUNIT: Forms homo- and heterologous complexes. Binds the TPR
 CC motif-containing C-terminal of kinesin light chain, Klc. Pre-
 CC assembled syd scaffolding complexes are then transported as a
 CC cargo of kinesin, to the required subcellular location.
 CC !- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC !- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=Q9QF1-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=Q9QF1-2; Sequence=VSP_002780;
 CC Note=No experimental confirmation available;
 CC !- SIMILARITY: Belongs to the JIP scaffold family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF262045; AAG36930.1; -;
 CC EMBL; AE003557; AAF50505.2; -;
 CC EMBL; AE003557; AAN12032.1; -;
 CC EMBL; AY060748; AAL28296.1; -;
 CC FlyBase; FBgn024187; syd.
 CC GO; GO:0005737; Cytoplasm; ISS.
 CC GO; GO:0030140; C:trans-Golgi network transport vesicle; IDA.
 CC GO; GO:0019894; F:kinesin binding; IPI.
 CC GO; GO:0005078; F:MAP-kinase scaffold activity; ISS.
 CC GO; GO:0019901; F:protein kinase binding; ISS.
 CC GO; GO:0008088; P:axon cargo transport; IMP.
 CC GO; GO:0046328; P:regulation of JNK cascade; ISS.
 CC GO; GO:0016192; P:vesicle-mediated transport; IEP.
 CC Coiled coil; Alternative splicing.
 CC DOMAIN 84 191 COILED COIL (POTENTIAL).
 CC FT DOMAIN 363 489 COILED COIL (POTENTIAL).
 CC FT DOMAIN 814 849 COILED COIL (POTENTIAL).
 CC FT VARSPLIC 1 55 Missing (in isoform Short).
 CC FT CONFLICT 983 983 /FTId=VSP_002780.
 CC FT SEQUENCE 1227 AA; 136711 MW; 2162C18066A057B8 CRC64;
 CC
 CC Query Match 35.1%; Score 1217.5; DB 1; Length 1227;
 CC Best Local Similarity 41.1%; Fred. NO. 8e-67;
 CC Matches 287; Conservative 70; Mismatches 174; Indels 167; Gaps 20;
 CC
 CC 4 VPVPVYCRPLVEKDPMTKMLCAAGVNLGSGWRPNEDDAGVKKPAPGRDPLTCDREGDGP 63
 CC 641 VPVPVYCNPLAEASPHMKVFCAGVNLHG-----GFTKNGQSLIPANSFYA-----P 687
 CC 64 KSA-----HTSPKKKAKEL-----PEMDATSSRVMLTTLTTSKWIID 104
 CC 688 KSTLKIAEITSPATQDSMEALDRQMARVSLTLEPQTQ-LSSFVWICTSTHAATSVSWD 746
 CC 105 ANQPGTVVDQTVCAHVLICSSIPASDSY-----PPGEMFLDSVNPEDPGAD 155
 CC 747 ANQATSLVDAPPI CASHLLCIASVQGAESDYALLEQSEVVKAGEML-----QRPQ-- 797
 CC 156 GVLAGITLVGCATCNVPRNSCSRSGDTPVLVDKGGEVATITANGKVPNSQSTEATEATE 215
 CC 798 ---EGTELLGKVEFVRV-----KPKSDDRQ-----NSNEKQQQEEBEAKATE 837
 CC 216 -----VDPDGPSEP-----ETATLR-----PGPLTEHVFTDPAFTPSGPGQSGENGPEPDSS 263

Db 838 KSEQLPAVSAEPELGNVEAIKIRQLPGA-----PQRUSTDGNQTNNNNNSSSN 889
 QY 264 -----STRPEPEPSGDPGTAGSSAAPTWMLGQNGWLVVHSAVANWKKLSIKLKOSV 317
 Db 890 LLFATKSLNPILETKDRDEPAMSSVGPMTWMLGAQDGLVYVHSSVGRWHECLHRVLLPDAV 949
 QY 318 LSLVHVKGRLVALADGTILAFHRGEDQWDLNSYHLMDLGHPHHSIRCMVAVYDRVWCG 377
 Db 950 LAIVHVEARVVVALANAQLAVERRQTDGQWDLNSYHLVTLGDRNHSIRCLCVAGERIWA 1009
 QY 378 YNKKVHVIOPTKMTOKESFSDAHPRESOVROLAWIGDVWVSIKRLDSTLRVYHATHOHL 437
 Db 1010 HRNKIFIVDPVSLNIVHSLDAHPKESQVRQAAATGAGVWVSIKRLDSTLRVYHATHOHL 1069
 QY 438 QVDIEPYVSKMLGTGKLGFSFVRITALLVAGSLWVGTVNGVWISIPLETVTWVLRGOL 497
 Db 1070 QVDIEPYVSKMLGTGKLGFSFVRITALLVAGSLWVGTVNGVWISIPLETVTWVLRGOL 1120
 QY 498 LGLRANKTSPTSGECAFGGIIHVYGDSSDRAASSFTPYCSMAQQLCFHGHRDAVKFF 557
 Db 1121 --VQPKSSSDPHGQ-----MPLCCMANAQLSFHGHRDAVKFF 1155
 QY 558 VSPGCVNLATLNGSVLDSPAESGPGPAPASEVEGOKLRNLVLSGGEGVIDPRIGDG-- 614
 Db 1156 VSPVLMQOPNLNGGLTFT-----NKRPRMLVMCGEGVIDFRINDNME 1199
 QY 615 -----EDDETEGAGDSQVKKPVLKASRSHIIVMQVS 647
 Db 1200 NSIQLEPNQTIENRGD-----KSYLIVWHVS 1225
 RESULT 4
 JIP-CABEL
 ID JIP CABEL STANDARD; PRT; 1157 AA.
 AC P34609; Q95V72;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE JNK-interacting protein (JIP) (JNK MAP kinase scaffold protein)
 DE (Uncoordinated protein 16).
 GN UNC-16 OR ZK1098.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCHI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM C), FUNCTION, INTERACTION WITH JNK-1;
 RP JKK-1 AND SEK-1, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
 RX MEDLINE=99321749; PubMed=10393177;
 RA Kawasaki M., Hisamoto N., Iino Y., Yamamoto M., Ninomiya-Tsuji J.,
 RA Matsumoto K.;
 RT "A Caenorhabditis elegans JNK signal transduction pathway regulates
 RT coordinated movement via type-D GABAergic motor neurons.";
 RL EMO J. 18:3604-3615(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
 RA Sims M., Smalton N., Smith A., Smith M., Sonnenhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).


```

CC  -!- SIMILARITY: Contains 1 DEL-homology (DH) domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; U00055; AAL06044.3; -
CC  WormPep; R02F2.2; CE31046.
CC  InterPro; IPR001331; GDS_CDC24.
CC  InterPro; IPR000219; RhoGEF.
CC  Pfam; PF00621; RhoGEF; 1.
CC  SMART; SM00325; RhoGEF; 1.
CC  PROSITE; PS00741; DH_1; 1.
CC  PROSITE; PS0010; DH_2; 1.
CC  KW  Hypothetical protein; Signal.
CC  FT  SIGNAL 1 15
CC  FT  CHAIN 16 1222
CC  FT  DOMAIN 83 118
CC  FT  DOMAIN 397 584
CC  FT  DOMAIN 1222 AA; 137354 MW; C12D723AF9D72973 CRC64;
CC  SQ  SEQUENCE 1222 AA; 137354 MW; C12D723AF9D72973 CRC64;
CC
CC  Query Match
CC  Best Local Similarity 5.3%; Score 183.5; DB 1; Length 1222;
CC  Matches 107; Conservative 80; Mismatches 193; Indels 144; Gaps 23;
CC
CC  QY 162 TLVGCATR-----CN-----VPRNSCSSRGDTPVL-----DKQGEVATIANGVNPSQSTE 208
CC  Db 805 TIIAHQTTPGLQLCTATVPGRVDS---TPSLWVCASDKFSGQVAVNA---LDTGETTI 858
CC
CC  QY 209 EATEA-----TEVPDPGPSEPETATLRPGPLFHEVFTDPAFTPSSGPQPSGNGPEPD 261
CC  Db 859 ESCSAIGNAAVTAMCTVPPMKLRKIKSKSLEHL-----NETIMD 902
CC
CC  QY 262 SSSTRPEPESGDPGTGASSAAPTMTLGAQNGWLYVHSAVANWKKCLHS--IKLKDSVLS 319
CC  Db 903 INSSGSDTESSD-EGTSTAGQTTWIGNDGGEVFNSTERSRARDRLARUNITS 961
CC
CC  QY 320 LVHVKGRLVALAGT---LAIHRGDEGQWDLNSYHMLDGHPHHS-IRCMAYVYDRVW 375
CC  Db 962 ICAANGNLVATSYSNQVLLFRPASDGSWLEN---PQTGVHVCQAPITSMQLIGRVI 1019
CC
CC  QY 376 CGYKNGVHVLPQKTMQLEKSFDAHPRESQVQLAWI-GDGVWTSIRLDSTLRLYHAHTH 434
CC  Db 1020 IASGNWLHAYFVDTGKFPPEVILP--SSDVITLMYVTGQVFLCGRKSTEVFVVDVFNL 1077
CC
CC  QY 435 QHLQDVVDIEPVSRMLG-----TGKLGFSFVRITALLVAGSLWVGTNGVVISIPLT 487
CC  Db 1078 SIINHFNVSFVRQLSGREHLREHKG--CLRISCLTVARSHLWGTSGACVLS----1131
CC
CC  QY 488 ETVVLHKGQLLGRANKTSPTSGEARPGGIIHVYGDSDRAASSFPYCSMAQAQLCF 547
CC  Db 1132 -----TSVQSARSQPT-----PDLRVCE 1149
CC
CC  QY 548 HGRDAVKFVSPGVNGLATINGSVLDSPAEGPGA-----APASEVSGKLRNVVLVSGG 603
CC  Db 1150 IGH-----SGPCRILLPVH-----TPSHNSHPSRKQRKSSLNVPAAQSSQLMLVSCG 1196
CC
CC  QY 604 EGYIDFRIGDGEDDETEGAGMSQVKPVLKSAERSHIIYVOVS 647
CC  Db 1197 EGLDD---GTATQDPSTDAI-----NHLIFWKS 1222
CC
CC  RESULT 6
CC  SHK3 HUMAN
CC  ID SHK3 HUMAN STANDARD; PRT; 797 AA.
CC  AC Q9BYE0; O8TET3;
CC  DT 28-FEB-2003 (Rel. 41, Created)
CC  DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC  DT 10-OCT-2003 (Rel. 42, Last annotation update)

```

```

DE SH3 and multiple ankyrin repeat domains protein 3 (Shank3) (Proline-
DE rich synapse-associated protein 2) (ProSAP2) (Fragment).
GN SHANK3 OR PSAP2 OR KIAA1650.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21156230; PubMed=11258795;
RA Hirosewa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.;
RT "Identification of novel transcribed sequences on human chromosome 22
RT by expressed sequence tag mapping.";
RL DNA Res. 8:1-9(2001).
RN [2]
RP SEQUENCE OF 28-797 FROM N.A.
RC TISSUE=Spleen;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHROMOSOMAL TRANSLOCATION.
RX MEDLINE=21344586; PubMed=11431708;
RA Bonaglia M.C., Giorda R., Borgatti R., Felisari G., Gagliardi C.,
RA Selicorni A., Zuffardi O.;
RT "Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is
RT associated with the 22q13.3 deletion syndrome.";
RL Am. J. Hum. Genet. 69:261-268(2001).
RN [4]
RP REVIEW.
RX MEDLINE=20267867; PubMed=10806096;
RA Sheng M., Kim E.;
RT "The Shank family of scaffold proteins.";
RL J. Cell Sci. 113:1851-1856(2000).
CC -!- FUNCTION: Seems to be an adapter protein in the postsynaptic
CC density (PSD) of excitatory synapses that interconnects receptors
CC of the postsynaptic membrane including NMDA-type and metabotropic
CC glutamate receptors via complexes with GKAP/PSD-95 and Homer,
CC respectively, and the actin-based cytoskeleton. May play a role in
CC the structural and functional organization of the dendritic spine
CC and synaptic junction.
CC -!- SUBUNIT: May homomultimerize via its SAM domain (By similarity).
CC Interacts with DLGAP1/GKAP, MGLUR1A, MGLUR5 C-termini via its PDZ
CC domain (By similarity). Interacts with Homer-1, Homer-2, Homer-3
CC and CTN/cortactin SH3 domain (By similarity). Is part of a
CC complex with DLG4/PSD-95 and DLGAP1/GKAP (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; postsynaptic density of
CC neuronal cells (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in the cerebral cortex and the
CC cerebellum.
CC -!- DISEASE: Involved in the chromosome 22q13.3 deletion syndrome by a
CC chromosomal translocation t(12;22)(q24.1;q13.3) that involves
CC D1P13B/FLJ10659 and SHANK3.
CC -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB051437; BAB33320.1; -
CC EMBL; AK074038; BAB84864.1; -
CC Genew; HGNC:14294; SHANK3.
CC MIM; 606230; -
CC MIM; 606232; -
CC InterPro; IPR001660; SAM.
CC Pfam; PF00536; SAM; 1.
CC SMART; SM00454; SAM; 1.
CC PROSITE; PS50105; SAM_DOMAIN; 1.

```

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; AJ133120; CAB45688.1; --
 DR EMBL; AF133301; AAF61375.1; ALT_FRAME.
 DR EMBL; AF159047; AAD42976.1; ALT_FRAME.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR001660; SAM.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00023; ank; 5.
 DR Pfam; PF00595; PDZ; 1.
 DR Pfam; PF00536; SAM; 1.
 DR Pfam; PF00018; SH3; 1.
 DR ProDom; PD000066; SH3; 1.
 DR SMART; SM00248; ANK; 5.
 DR SMART; SM00228; PDZ; 1.
 DR SMART; SM00454; SAM; 1.
 DR SMART; SM00326; SH3; 1.
 DR PROSITE; PS50088; ANK REPEAT; 4.
 DR PROSITE; PS50297; ANK REP REGION; 1.
 DR PROSITE; PS50106; PDZ; 1.
 DR PROSITE; PS50105; SAM DOMAIN; 1.
 DR PROSITE; PS50002; SH3; 1.
 KW ANK repeat; SH3 domain; SH3-binding; Coiled coil; Repeat;
 KW Alternative splicing.
 FT REPEAT 223 256 ANK 1.
 FT REPEAT 257 289 ANK 2.
 FT REPEAT 324 356 ANK 3.
 FT REPEAT 357 389 ANK 4.
 FT DOMAIN 472 488 COILED COIL (POTENTIAL).
 FT DOMAIN 545 604 SH3.
 FT DOMAIN 645 739 SH3-BINDING (POTENTIAL).
 FT SITE 1486 1492 SH3-BINDING (POTENTIAL).
 FT DOMAIN 1570 1590 COILED COIL (POTENTIAL).
 FT DOMAIN 1752 1815 SAM.
 FT DOMAIN 6 15 POLY-ALA.
 FT DOMAIN 895 901 POLY-PRO.
 FT DOMAIN 914 919 POLY-PRO.
 FT VARSPLIC 1204 1211 SPTVHPSP -> PRRRAGMV (in isoform 3).
 FT VARSPLIC 1212 1815 /FTid=VSP 006087.
 FT REPEAT 1612 1620 /FTid=VSP 006088.
 FT VARSPLIC 1386 1386 Missing (in isoform 1).
 FT MUTAGEN 1389 P->L: ABOLISHES INTERACTION WITH HOMER-1
 FT MUTAGEN 1389 F->C: ABOLISHES INTERACTION WITH HOMER-1
 FT CONFLICT 1 2 ISOFORM C.
 FT CONFLICT 32 32 MQ -> AA (IN REF. 2).
 FT CONFLICT 37 37 R -> Y (IN REF. 2).
 FT CONFLICT 68 68 P -> Q (IN REF. 2).
 FT CONFLICT 472 472 Q -> P (IN REF. 2).
 FT CONFLICT 781 781 H -> L (IN REF. 3).
 FT CONFLICT 902 907 R -> G (IN REF. 3).
 FT CONFLICT 912 917 YFDSG -> ILRLR (IN REF. 3).
 FT CONFLICT 921 921 FSPPPP -> SHHGHQ (IN REF. 3).
 FT CONFLICT 963 963 R -> G (IN REF. 3).
 FT CONFLICT 1115 1115 S -> N (IN REF. 2).
 FT CONFLICT 1162 1162 G -> S (IN REF. 2).
 FT CONFLICT 1167 1167 S -> N (IN REF. 2).
 FT CONFLICT 1337 1337 S -> N (IN REF. 2).
 FT CONFLICT 1345 1345 G -> S (IN REF. 2).
 FT CONFLICT 1348 1348 S -> N (IN REF. 2).
 FT CONFLICT 1354 1354 E -> K (IN REF. 2).
 FT CONFLICT 1369 1369 S -> N (IN REF. 2).

FT CONFLICT 1507 1507 D -> G (IN REF. 2).
 FT CONFLICT 1715 1715 V -> A (IN REF. 2).
 SQ SEQUENCE 1815 AA; 193256 MW; 6632CF8F040766F9 CRC64;
 Query Match 4.2%; Score 146.5; DB 1; Length 1815;
 Best Local Similarity 21.3%; Pred No. 0.31;
 Matches 147; Conservative 83; Mismatches 267; Indels 193; Gaps 30;
 QY 37 EDDAGNGVKA-----PCRDPLTCDREGDGEKPSAHTSPEKKKAKELPMDATSSR----- 87
 DB 1227 DSEKGLASPAFSPRSPAWIPVAPARREAEKPTREERKSPEDKSMLSVLDTSLQRPAGU 1286
 QY 88 -VMILTSTLTTSKVIIIDANQGTVDQFTVCNAHVLCISSIPAAASDSDYPGEMFLDSD 146
 DB 1287 IVVHATSNQBNRLGAEERPGTPELATPMQAAAV---AEPMPSPRAQPPGSI----- 1338
 QY 147 VNPEDPG-----ADGVLAGITLVGCATRCNVPNSCRSGDTPLVDKQCGGVANTIA 197
 DB 1339 --PADPGQGSSSEEPVLVA-----VNLPPAQLSSD-----EETREELARI- 1380
 QY 198 NGKNPSQSTEEATEATVPDPGPSEPTATLRPGLTTEHVFTDPAPTSSGPQGSSENG 257
 DB 1381 -GLVPP---PEEFANGILLATPPPG-----PGPLPTTV-----PSPASG-KPSSELP 1422
 QY 258 PEPDSSS-----TRPEPEPGDPTGAGSAAPTMGLQNGWLY-VHSAVANWKKCL 308
 DB 1423 PAPSAADSGVEEADTRSSPHLETTTITVSSMSTLSESGELTDTHTSFADG----- 1478
 QY 309 HSIKIKDSVLSLVHVKGRVLVALADGTLA----IFHRGEDGQWDLNHYLMDLGHPHHSI 364
 DB 1479 HTFLEKEPP---VPPKPKLKSPLGKGPVTFRDPLLKQSSDSEL-----MAQOHR- 1525
 QY 365 RCMAVYDVRVWCGYKKNVHVITQPKIMQIEKSFDAHPRESQVRLAWIGDGV----- 416
 DB 1526 -----TSTGLTSAAGPARPYLFQRRSKLW-GDPVESRGLPGP 1562
 QY 417 -----WYSIRLDSTLRVYHANTHOHQVDIEPVYKMLGTGKLGFSFVRITALLVAGSR 471
 DB 1563 EDDKPTVISELSSRLQINKDT---RSLGEEP-----VGLGLSLDPAKSPIAAAR 1611
 QY 472 LWVGTVNGVVISIPLTETVHLRGQLGLLRANKTSPTSGEGA---RPGGIHHVVGDDSS 527
 DB 1612 CAV-----VPSAGNLSSLGELSTISNQRSPGPGGASVSRSPGYPV----- 1656
 QY 528 DRAASSIPYCSMAQAOLCFHGHDAVKFFVSPGVNVLATNGSV----- 572
 DB 1657 ARRAPSVPKPSALRVEGLGAGVGAGRPGLTPTTILKSSLSIPIHEPKVRFVVRVS 1716
 QY 573 -----LDSPAEGGPAA-PASEVEGQKLRLNVLVSGEGYIDFRIGDGED---DE 618
 DB 1717 ARSRSPSPSLPSPSGSGPSAGRRPFQKPLQLWSKFDVGDWLESIHLEHGRDRFEDH 1776
 QY 619 TEEGAGDMSQVKPVLKSAERSHIIVQVSY 648
 DB 1777 EIEGAH-----LPALTKEDFVELGVTRVGH 1801
 RESULT 8
 AMYH YEAST
 ID AMYH YEAST STANDARD; PRT; 1367 AA.
 AC P08640; P08068;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glucosylase S1/S2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-
 DE Glucosidase) (1,4-alpha-D-glucan glucosylase).
 GN STA1 OR STA2 OR MAL5 OR YIR019C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;
RX MEDLINE=97313266; PubMed=9169870;
RA Churche C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Iye G.,
RA Maule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.;"
RL Nature 387:84-87(1997).
RN [2]
RP SEQUENCE OF 1-242 AND 762-1331 FROM N.A.
RX MEDLINE=87194600; PubMed=3106330;
RA Yamashita I., Nakamura M., Fukui S.;
RT "Gene fusion is a possible mechanism underlying the evolution of
RT STAL.;"
RL [3]
RN J. Bacteriol. 169:2142-2149 (1987).
RP SEQUENCE OF 1-31 FROM N.A.
RX STRAIN=SPX101-1C; PubMed=3141213;
RA Pardo J.M., Ianez E., Zalacain M., Claros M.G., Jimenez A.;
RT "Similar short elements in the 5' regions of the STA2 and SGA genes
RT from Saccharomyces cerevisiae.;"
RL FEBS Lett. 239:179-184(1988).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
CC glucose residues successively from non-reducing ends of the chains
CC with release of beta-D-glucose.
CC -!- SIMILARITY: TO S.POMBE SPC215.13.
CC -!- SIMILARITY: SOME, TO S.POMBE SPC285.13C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; Z38061; CAA86176.1; -;
DR EMBL; M16164; AAA35014.1; -;
DR EMBL; M16165; AAA35015.1; -;
DR EMBL; X13857; CAA32069.1; -;
DR PIR; S48478; S48478.
DR GeneOnline; 139731; -;
DR SGD; S000458; MUC1.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0030447; P:filamentous growth; IDA.
DR GO; GO:0007125; P:invasive growth; IMP.
DR GO; GO:0007124; P:pseudohyphal growth; IMP.
KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
KW Signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 1367
FT DOMAIN 210 1367
FT SER/THR-RICH.
FT CARBOHYD 817 817
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 874 874
FT N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1367 AA; 136110 MW; 91C00E2BDB61AA9D CRC64;

Query Match 4.0%; Score 138; DB 1; Length 1367;
Best Local Similarity 24.2%; Pred. No. 0.73;
Matches 60; Conservative 25; Mismatches 121; Indels 42; Gaps 8;

QY 47 AGRPLICDRGDEPKSAHTSPKKAKEPEMDATSSRWILTLTTSKVVILDAAN 106
DB 456 APVPTSSSTSSAPVTSSTSSAPVPTSSSTSSAPVTSSTSSAPVPTP 515
QY 107 QPGTVMDQFTVCNAHVLICISIPAAASDSYPPGEMFLDSVDNPDGADGLVIGITLVC 166
DB 516 SSSSTSS-----SSAPVPT-----PSSSTSSAP-----VTSSTSS 550
QY 167 ATRCNVPRNCSSRGTPV-----LDKGQGEVATIANGNVPSQSTEEATEVDPGGS 222
DB 551 SAPVPTSSSTSSSTSSPTVTSSTSSAPVPT-----PSSSTSSSA-PVPTPSSS 602

QY 223 EPETATL-RGCLTEHVFTDPAPTSSGPOGSENGPEPSSSTR-----PEPEPSGDPTG 277
DB 603 TTESSAPAPTPSSSTSSAPVTSSTSSAPVPTSSSTSSAPVPTSSSTSSAPVPTSSSTTE 662
QY 278 AGSSAAPT 285
DB 663 SSSAPVPT 670

RESULT 9
ID HWPI CANAL STANDARD; PRT; 634 AA.
AC P46593; O13424; P87019;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hyphal wall protein 1 (Cell elongation protein 2).
GN HWPI OR ECE2.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SC5314;
RX MEDLINE=96198091; PubMed=8626424;
RA Staab J.F., Ferrer C.A., Sundstrom P.;
RT "Developmental expression of a tandemly repeated, proline-and
RT glutamine-rich amino acid motif on hyphal surfaces on Candida
RT albicans.;"
RL J. Biol. Chem. 271:6298-6305(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SC5314;
RA Sharkey L.L., Saporito-Irwin S.M., Fonzi W.A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-225 FROM N.A.
RC STRAIN=SC5314;
RA Staab J.F., Ferrer C.A., Sundstrom P.R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Hyphal surface antigen.
CC -!- SUBCELLULAR LOCATION: Hyphal surface.
CC -!- TISSUE SPECIFICITY: Found in hyphal but not yeast forms.
CC -!- PTM: O-glycosylated (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U64206; AAC96368.1; -;
DR EMBL; AF001978; AAB64014.1; -;
DR EMBL; U29369; AAC49209.1; -;
KW Glycoprotein; Repeat; Antigen.
FT DOMAIN 46 187
FT REPEAT 46 58
FT REPEAT 59 69
FT REPEAT 70 81
FT REPEAT 82 91
FT REPEAT 92 101
FT REPEAT 102 111
FT REPEAT 112 121
FT REPEAT 122 131
FT REPEAT 132 141
FT REPEAT 142 151
FT REPEAT 152 161
FT REPEAT 162 171

14 X 10 AA TANDEM REPEATS OF [EVIQ]-P-
[CDT]-D-[YNW]-P-[PO]-[QI]-[QP]-[QDN].
1 (APPROXIMATE).
2 (APPROXIMATE).
3 (APPROXIMATE).
4.
5.
6.
7.
8.
9.
10.
11.
12.

```

FT REPEAT 172 179 13 (INCOMPLETE) .
FT REPEAT 180 187 14 (INCOMPLETE) .
FT DOMAIN 53 58 POLY-GLN.
FT DOMAIN 65 69 POLY-GLN.
FT DOMAIN 204 207 POLY-THR.
FT DOMAIN 208 220 POLY-SER.
FT DOMAIN 291 296 POLY-THR.
FT DOMAIN 304 307 POLY-THR.
FT DOMAIN 398 401 POLY-VAL.
FT DOMAIN 141 141 D -> DVPCDNPQPD (IN REF. 3) .
FT CONFLICT 296 296 T -> A (IN REF. 2) .
FT CONFLICT 441 441 S -> P (IN REF. 2) .
FT CONFLICT 492 492 P -> S (IN REF. 2) .
SQ SEQUENCE 634 AA; 653372 MW; F841347576BA6376 CRC64;

Query Local
Best Local Similarity 23.8%; Score 137.5; DB 1; Length 634;
Matches 70; Conservative 35; Mismatches 98; Indels 91; Gaps 14;

QY 37 EDDAGNGVKPAPGRD-----PLTCDREGDGPKSATSPKPKKAKELPE---MDATSS 86
Db 326 ESEVITGVIVITSKDIIVTYICPLT-----ETTPVSTAP----ATETPTGTVTSTEQ 374
QY 87 RWIILT-----SLTITSKVI-----IDANQPGTVVDQFTVCNAHVL 123
Db 375 STTIVITVSCBSCTSESVITGVVVTSEETVYITFCPLTENTFGT----DSTPEASIP 430
QY 124 CISSIPAAASDSDYRPGEMFLDSVDNPDGADGVLGITLVCCATRCNVPRNCSRSRGDT 183
Db 431 PMETIPAGESMPAGE-----TSPAVPKSDVPATESA 463
QY 184 PVLID-KGQGEVATIANGVNPS--QSTBEATEATEVDPDGPSEPTATLRP---GPLTEH 237
Db 464 PVPEMTAGSQSPISAGETSPAVPKSDVPATESAPAPMTAGTITKAAPKSSAPATB- 522
QY 238 VETDAP-----TPSSGPOPGSENGEPDSSSTRPEP-EPSGDPTGAGSSAAPT 286
Db 523 ----PSPVAPGTESAPAGGAGSSPKSVLASSETSPAPGAEATAPAGSSGAIIT 572

RESULT 10
MINT_MOUSE
ID MINT_MOUSE STANDARD; PRT; 3644 AA.
AC Q62504; Q80TN9; Q99PS4; Q9QZW2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Mx2-interacting protein (SMART/HDAC1 associated repressor protein) .
GN MINT OR SHARP OR KIAA0929.
OS Mus musculus (Mouse) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE OF 49-3644 FROM N.A. (ISOFORM 1), FUNCTION, SUBCELLULAR
RP LOCATION, TISSUE SPECIFICITY, DNA-BINDING, AND INTERACTION WITH MSX2.
RC TISSUE=Testis;
RX MEDLINE=93379811; PubMed=10451362;
RA Newberry E.P., Latifi T., Towler D.A.;
RT "The RRM domain of MINT, a novel mxx2 binding protein, recognizes and
RL regulates the rat osteocalcin promoter.";
RL Biochemistry 38:10678-10690(1999).
RN [2]
RP SEQUENCE OF 1-112 FROM N.A.
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S.,
RA Hashizume W., Hayashida K., Hirozane T., Hori F., Imotani K.,
RA Ishii Y., Itoh M., Kagawa I., Kawai J., Kojima Y., Kondo S., Konno H.,
RA Koya S., Miyazaki A., Murata M., Nakamura M., Nomura K., Numazaki R.,
RA Ohno K., Ohsato N., Saito R., Sakazume N., Sano H., Sasaki D.,
RA Sato K., Shibata K., Shiraki T., Tagami M., Takeda Y., Waki K.,
RA Watahiki A., Muramatsu M., Hayashizaki Y.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [3]

```

```

RP SEQUENCE OF 69-3644 FROM N.A. (ISOFORM 2), AND VARIANTS THR-348;
RP PHE-762; PHE-773 AND LEU-933.
RC STRAIN=ICR; TISSUE=Brain;
RA Sakamoto T., Gotou T., Isagawa Y., Mimura H., Kimura K., Kawaichi M.;
RT "MINT/spen negatively regulates Notch signaling by inhibiting RBP-
RT J/Su(H) activity.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 318-578 FROM N.A.
RP TISSUE=Cochlea;
RX MEDLINE=97237053; PubMed=9119401;
RA Crozet F., El-Amrani A., Blanchard S., Lenoir M., Ripoll C., Vago P.,
RA Hamel C., Fizames C., Levi-Acobas F., Depetris D., Mattei M.-G.,
RA Weil D., Pujol R., Petit C.;
RT "Cloning of the genes encoding two murine and human cochlear
RT unconventional type I myosins.";
RL Genomics 40:332-341(1997).
RN [5]
RP SEQUENCE OF 2598-3644 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RL randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
RN [6]
RP TISSUE SPECIFICITY.
RX MEDLINE=22261914; PubMed=12374742;
RA Oswald F., Kosterka U., Astrahantseff K., Bourteele S., Dillinger K.,
RA Zechner U., Ludwig L., Wilda M., Hameister H., Knoechel W., Liptay S.,
RA Schmid R.M.;
RT "SHARP is a novel component of the Notch/RBP-Jkappa signalling
RT pathway.";
RL EMBO J. 21:5417-5426(2002).
RN [7]
RP FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=22483652; PubMed=12594956;
RA Kuroda K., Han H., Tani S., Tanigaki K., Tun T., Furukawa T.,
RA Taniguchi Y., Kurooka H., Hamada Y., Toyokuni S., Honjo T.;
RT "Regulation of marginal zone B cell development by MINT, a suppressor
RT of Notch/RBP-J signaling pathway.";
RL Immunity 18:301-312(2003).
CC -!- FUNCTION: Essential corepressor protein, which probably regulates
CC different key pathways such as the Notch pathway. Negative
CC regulator of the Notch pathway via its interaction with RBPSUH,
CC which prevents the association between NOTCH1 and RBPSUH, and
CC therefore suppresses the transcription activity of Notch
CC signaling. Blocks the differentiation of precursor B cells into
CC marginal zone B cells. Probably represses transcription via the
CC recruitment of large complexes containing histone deacetylase
CC proteins. May bind both to DNA and RNA.
CC -!- SUBUNIT: Interacts with NCOR2, HDAC1, HDAC2, RBBP4, MBD3 and
CC MYALL. Interacts with the nuclear receptors RAR and PPARG.
CC Interacts with RAR in absence of ligand. Bind to the steroid
CC receptor RNA coactivator SRA (By similarity). Interacts with MSX2.
CC Interacts with RBPSUH; this interaction may prevent the
CC interaction between RBPSUH and NOTCH1.
CC -!- SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q62504-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q62504-2; Sequence=VSP_008564;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Highly expressed in testis. Expressed at lower
CC level in brain, lung, spleen, liver and kidney. Weakly expressed
CC in cardiac and skeletal muscles and ovary. In spleen, it is
CC expressed in follicular B-cells, while it is weakly expressed in
CC marginal zone B-cells.

```


OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Placenta;
RX MEDLINE=20143540; PubMed=10677613;
RA Zagon I.S., Verderame M.F., Allen S.S., McLaughlin P.J.;
RT "Cloning, sequencing, chromosomal location, and function of cDNAs
RT encoding an opioid growth factor receptor (OGFr) in humans.";
RL Brain Res. 856:75-83(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Takanosu M., Liu J., Mayne R., Wood B.M., Brewton R.G.;
RT "Genomic structure of human gene 7-60";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levenshain M.H., Levenshain M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McManis L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuce C.D., Smith M.I., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin R., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20";
RN Nature 414:865-871(2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 37-677 FROM N.A. (ISOFORM 1).
RC TISSUE=Spleen;
RX MEDLINE=21082933; PubMed=11214971;
RA Hattori A., Okumura K., Nagase T., Kikuno R., Hirose M., Ohara O.;
RT "Characterization of long cDNA clones from human adult spleen.";
RN DNA Res. 7:357-366(2000).
RN [6]
RP REVIEW
RX MEDLINE=2188341; PubMed=11890982;
RA Zagon I.S., Verderame M.F., McLaughlin P.J.;
RT "The biology of the opioid growth factor receptor (OGFr).";
RL Brain Res. 38:351-376(2002).
CC -!- FUNCTION: Receptor for opioid growth factor (OGF), also known as
CC Met-enkephalin. Seems to be involved in growth regulation.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity). The OGFr/OGFR
CC complex is probably translocated to the nucleus.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9NZT2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NZT2-2; Sequence=VSP_004060;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Highly expressed in the heart and liver,
CC moderately in skeletal muscle and kidney and to a lesser extent in
CC brain and pancreas. Expressed in fetal tissues including liver and
CC kidney.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts
CC in position 1 and 181.
CC -!- CAUTION: Ref.3 (CAC28882) sequence differs from that shown due to
CC erroneous gene model prediction.
CC -!- CAUTION: Ref.3 (CAC12749) sequence differs from that shown due to
CC a frameshift in position 514.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF172451; AAF64404.1; -
CC EMBL; AF172452; AAF64405.1; -
CC EMBL; AF172453; AAF64406.1; -
CC EMBL; AF112980; AAD03745.1; ALT_FRAME.
CC EMBL; AF109134; AAD03737.1; ALT_FRAME.
CC EMBL; AF035669; CAC28882.1; ALT_SEQ.
CC EMBL; AL035669; CAC12749.1; ALT_FRAME.
CC EMBL; BC014137; AAH14137.1; -
CC EMBL; AK024485; BAB15775.1; -
CC Genew; HGNC:15768; OGFR.
CC MIM; 606459; -
CC GO; GO:0004985; F:opioid receptor activity; NAS.
CC GO; GO:0001558; P:regulation of cell growth; NAS.
CC InterPro; IPR006757; OGFR_N.
CC InterPro; IPR006770; OGFR_repeat.
CC Pfam; PF04664; OGFR_N; 1.
CC Pfam; PF04680; OGFR_repeat; 7.
CC Receptor; Growth regulation; Repeat; Alternative splicing.
CC DOMAIN 283 BIPARTITE NUCLEAR LOCALIZATION SIGNAL
CC (POTENTIAL).
CC 7 X APPROXIMATE 20 AA TANDEM REPEATS.
FT DOMAIN 517 656
FT REPEAT 517 536 1.
FT REPEAT 537 556 2.
FT REPEAT 557 576 3.
FT REPEAT 577 596 4.
FT REPEAT 597 616 5.
FT REPEAT 617 636 6.
FT REPEAT 637 656 7.
FT VARSPLIC 558 577 Missing (in isoform 2).
FT CONFLICT 27 28 /FTID=VSP_004060.
FT CONFLICT 37 80 DG -> EA (IN REF. 2).
FT DAGDEDESEEPRAAPSFQSRMTGSRNWRATRDRCYRH

RT subfamily, is highly expressed in the developing and adult central nervous system.";
RL Mech. Dev. 93:221-231(2000).
CC -!- TISSUE SPECIFICITY: Expressed almost exclusively in the developing and mature central nervous system.
CC -!- DEVELOPMENTAL STAGE: Expression begins at embryonic day 9.5 in the developing spinal cord and brain structures and continues in neonatal and adult brain structures including the olfactory bulb, cortex, thalamus, hypothalamus, hippocampus and cerebellum.
CC -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
CC -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
CC -!- SIMILARITY: Belongs to the formin homology family. Cappuccino subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF218940; AAF72883.1; -
CC MGD; MGI:1859252; Fnm2.
CC InterPro; IPR003104; FH2.
CC Pfam; PF02181; FH2; 1.
CC PRINTS; PRO0828; FORMIN.
CC SMART; SM00498; FH2; 1.
CC Developmental protein; Repeat; Coiled coil.
CC COILED COIL (POTENTIAL).
CC FH1 (PRO-RICH).
CC 11 X 11 AA TANDEM REPEATS OF [MV]-G-I-P-P-P-P-P-L-P-G.
CC P-P-P-P-P-L-P-G.
CC 1.
CC REPEAT 919 929
CC REPEAT 930 940
CC REPEAT 941 951
CC REPEAT 952 962
CC REPEAT 963 973
CC REPEAT 974 984
CC REPEAT 985 995
CC REPEAT 996 1006
CC REPEAT 1007 1017
CC REPEAT 1018 1028
CC REPEAT 1029 1039
CC REPEAT 1040 1050
CC REPEAT 1051 1061
CC REPEAT 1062 1072
CC REPEAT 1073 1083
CC REPEAT 1084 1094
CC REPEAT 1095 1105
CC REPEAT 1106 1116
CC REPEAT 1117 1127
CC REPEAT 1128 1138
CC REPEAT 1139 1149
CC REPEAT 1150 1160
CC REPEAT 1161 1171
CC REPEAT 1172 1182
CC REPEAT 1183 1193
CC REPEAT 1194 1204
CC REPEAT 1205 1215
CC REPEAT 1216 1226
CC REPEAT 1227 1237
CC REPEAT 1238 1248
CC REPEAT 1249 1259
CC REPEAT 1260 1270
CC REPEAT 1271 1281
CC REPEAT 1282 1292
CC REPEAT 1293 1303
CC REPEAT 1304 1314
CC REPEAT 1315 1325
CC REPEAT 1326 1336
CC REPEAT 1337 1347
CC REPEAT 1348 1358
CC REPEAT 1359 1369
CC REPEAT 1370 1380
CC REPEAT 1381 1391
CC REPEAT 1392 1402
CC REPEAT 1403 1413
CC REPEAT 1414 1424
CC REPEAT 1425 1435
CC REPEAT 1436 1446
CC REPEAT 1447 1457
CC REPEAT 1458 1468
CC REPEAT 1469 1479
CC REPEAT 1480 1490
CC REPEAT 1491 1501
CC REPEAT 1502 1512
CC REPEAT 1513 1523
CC REPEAT 1524 1534
CC REPEAT 1535 1545
CC REPEAT 1546 1556
CC REPEAT 1557 1567
CC REPEAT 1568 1578
CC REPEAT 1579 1589
CC REPEAT 1590 1599
CC REPEAT 1600 1610
CC REPEAT 1611 1621
CC REPEAT 1622 1632
CC REPEAT 1633 1643
CC REPEAT 1644 1654
CC REPEAT 1655 1665
CC REPEAT 1666 1676
CC REPEAT 1677 1687
CC REPEAT 1688 1698
CC REPEAT 1699 1709
CC REPEAT 1710 1720
CC REPEAT 1721 1731
CC REPEAT 1732 1742
CC REPEAT 1743 1753
CC REPEAT 1754 1764
CC REPEAT 1765 1775
CC REPEAT 1776 1786
CC REPEAT 1787 1797
CC REPEAT 1798 1808
CC REPEAT 1809 1819
CC REPEAT 1820 1830
CC REPEAT 1831 1841
CC REPEAT 1842 1852
CC REPEAT 1853 1863
CC REPEAT 1864 1874
CC REPEAT 1875 1885
CC REPEAT 1886 1896
CC REPEAT 1897 1907
CC REPEAT 1908 1918
CC REPEAT 1919 1929
CC REPEAT 1930 1940
CC REPEAT 1941 1951
CC REPEAT 1952 1962
CC REPEAT 1963 1973
CC REPEAT 1974 1984
CC REPEAT 1985 1995
CC REPEAT 1996 2006
CC REPEAT 2007 2017
CC REPEAT 2018 2028
CC REPEAT 2029 2039
CC REPEAT 2040 2050
CC REPEAT 2051 2061
CC REPEAT 2062 2072
CC REPEAT 2073 2083
CC REPEAT 2084 2094
CC REPEAT 2095 2105
CC REPEAT 2106 2116
CC REPEAT 2117 2127
CC REPEAT 2128 2138
CC REPEAT 2139 2149
CC REPEAT 2150 2160
CC REPEAT 2161 2171
CC REPEAT 2172 2182
CC REPEAT 2183 2193
CC REPEAT 2194 2204
CC REPEAT 2205 2215
CC REPEAT 2216 2226
CC REPEAT 2227 2237
CC REPEAT 2238 2248
CC REPEAT 2249 2259
CC REPEAT 2260 2270
CC REPEAT 2271 2281
CC REPEAT 2282 2292
CC REPEAT 2293 2303
CC REPEAT 2304 2314
CC REPEAT 2315 2325
CC REPEAT 2326 2336
CC REPEAT 2337 2347
CC REPEAT 2348 2358
CC REPEAT 2359 2369
CC REPEAT 2370 2380
CC REPEAT 2381 2391
CC REPEAT 2392 2402
CC REPEAT 2403 2413
CC REPEAT 2414 2424
CC REPEAT 2425 2435
CC REPEAT 2436 2446
CC REPEAT 2447 2457
CC REPEAT 2458 2468
CC REPEAT 2469 2479
CC REPEAT 2480 2490
CC REPEAT 2491 2501
CC REPEAT 2502 2512
CC REPEAT 2513 2523
CC REPEAT 2524 2534
CC REPEAT 2535 2545
CC REPEAT 2546 2556
CC REPEAT 2557 2567
CC REPEAT 2568 2578
CC REPEAT 2579 2589
CC REPEAT 2590 2599
CC REPEAT 2600 2610
CC REPEAT 2611 2621
CC REPEAT 2622 2632
CC REPEAT 2633 2643
CC REPEAT 2644 2654
CC REPEAT 2655 2665
CC REPEAT 2666 2676
CC REPEAT 2677 2687
CC REPEAT 2688 2698
CC REPEAT 2699 2709
CC REPEAT 2710 2720
CC REPEAT 2721 2731
CC REPEAT 2732 2742
CC REPEAT 2743 2753
CC REPEAT 2754 2764
CC REPEAT 2765 2775
CC REPEAT 2776 2786
CC REPEAT 2787 2797
CC REPEAT 2798 2808
CC REPEAT 2809 2819
CC REPEAT 2820 2830
CC REPEAT 2831 2841
CC REPEAT 2842 2852
CC REPEAT 2853 2863
CC REPEAT 2864 2874
CC REPEAT 2875 2885
CC REPEAT 2886 2896
CC REPEAT 2897 2907
CC REPEAT 2908 2918
CC REPEAT 2919 2929
CC REPEAT 2930 2940
CC REPEAT 2941 2951
CC REPEAT 2952 2962
CC REPEAT 2963 2973
CC REPEAT 2974 2984
CC REPEAT 2985 2995
CC REPEAT 2996 3006
CC REPEAT 3007 3017
CC REPEAT 3018 3028
CC REPEAT 3029 3039
CC REPEAT 3040 3050
CC REPEAT 3051 3061
CC REPEAT 3062 3072
CC REPEAT 3073 3083
CC REPEAT 3084 3094
CC REPEAT 3095 3105
CC REPEAT 3106 3116
CC REPEAT 3117 3127
CC REPEAT 3128 3138
CC REPEAT 3139 3149
CC REPEAT 3150 3160
CC REPEAT 3161 3171
CC REPEAT 3172 3182
CC REPEAT 3183 3193
CC REPEAT 3194 3204
CC REPEAT 3205 3215
CC REPEAT 3216 3226
CC REPEAT 3227 3237
CC REPEAT 3238 3248
CC REPEAT 3249 3259
CC REPEAT 3260 3270
CC REPEAT 3271 3281
CC REPEAT 3282 3292
CC REPEAT 3293 3303
CC REPEAT 3304 3314
CC REPEAT 3315 3325
CC REPEAT 3326 3336
CC REPEAT 3337 3347
CC REPEAT 3348 3358
CC REPEAT 3359 3369
CC REPEAT 3370 3380
CC REPEAT 3381 3391
CC REPEAT 3392 3402
CC REPEAT 3403 3413
CC REPEAT 3414 3424
CC REPEAT 3425 3435
CC REPEAT 3436 3446
CC REPEAT 3447 3457
CC REPEAT 3458 3468
CC REPEAT 3469 3479
CC REPEAT 3480 3490
CC REPEAT 3491 3501
CC REPEAT 3502 3512
CC REPEAT 3513 3523
CC REPEAT 3524 3534
CC REPEAT 3535 3545
CC REPEAT 3546 3556
CC REPEAT 3557 3567
CC REPEAT 3568 3578
CC REPEAT 3579 3589
CC REPEAT 3590 3599
CC REPEAT 3600 3610
CC REPEAT 3611 3621
CC REPEAT 3622 3632
CC REPEAT 3633 3643
CC REPEAT 3644 3654
CC REPEAT 3655 3665
CC REPEAT 3666 3676
CC REPEAT 3677 3687
CC REPEAT 3688 3698
CC REPEAT 3699 3709
CC REPEAT 3710 3720
CC REPEAT 3721 3731
CC REPEAT 3732 3742
CC REPEAT 3743 3753
CC REPEAT 3754 3764
CC REPEAT 3765 3775
CC REPEAT 3776 3786
CC REPEAT 3787 3797
CC REPEAT 3798 3808
CC REPEAT 3809 3819
CC REPEAT 3820 3830
CC REPEAT 3831 3841
CC REPEAT 3842 3852
CC REPEAT 3853 3863
CC REPEAT 3864 3874
CC REPEAT 3875 3885
CC REPEAT 3886 3896
CC REPEAT 3897 3907
CC REPEAT 3908 3918
CC REPEAT 3919 3929
CC REPEAT 3930 3940
CC REPEAT 3941 3951
CC REPEAT 3952 3962
CC REPEAT 3963 3973
CC REPEAT 3974 3984
CC REPEAT 3985 3995
CC REPEAT 3996 4006
CC REPEAT 4007 4017
CC REPEAT 4018 4028
CC REPEAT 4029 4039
CC REPEAT 4040 4050
CC REPEAT 4051 4061
CC REPEAT 4062 4072
CC REPEAT 4073 4083
CC REPEAT 4084 4094
CC REPEAT 4095 4105
CC REPEAT 4106 4116
CC REPEAT 4117 4127
CC REPEAT 4128 4138
CC REPEAT 4139 4149
CC REPEAT 4150 4160
CC REPEAT 4161 4171
CC REPEAT 4172 4182
CC REPEAT 4183 4193
CC REPEAT 4194 4204
CC REPEAT 4205 4215
CC REPEAT 4216 4226
CC REPEAT 4227 4237
CC REPEAT 4238 4248
CC REPEAT 4249 4259
CC REPEAT 4260 4270
CC REPEAT 4271 4281
CC REPEAT 4282 4292
CC REPEAT 4293 4303
CC REPEAT 4304 4314
CC REPEAT 4315 4325
CC REPEAT 4326 4336
CC REPEAT 4337 4347
CC REPEAT 4348 4358
CC REPEAT 4359 4369
CC REPEAT 4370 4380
CC REPEAT 4381 4391
CC REPEAT 4392 4402
CC REPEAT 4403 4413
CC REPEAT 4414 4424
CC REPEAT 4425 4435
CC REPEAT 4436 4446
CC REPEAT 4447 4457
CC REPEAT 4458 4468
CC REPEAT 4469 4479
CC REPEAT 4480 4490
CC REPEAT 4491 4501
CC REPEAT 4502 4512
CC REPEAT 4513 4523
CC REPEAT 4524 4534
CC REPEAT 4535 4545
CC REPEAT 4546 4556
CC REPEAT 4557 4567
CC REPEAT 4568 4578
CC REPEAT 4579 4589
CC REPEAT 4590 4599
CC REPEAT 4600 4610
CC REPEAT 4611 4621
CC REPEAT 4622 4632
CC REPEAT 4633 4643
CC REPEAT 4644 4654
CC REPEAT 4655 4665
CC REPEAT 4666 4676
CC REPEAT 4677 4687
CC REPEAT 4688 4698
CC REPEAT 4699 4709
CC REPEAT 4710 4720
CC REPEAT 4721 4731
CC REPEAT 4732 4742
CC REPEAT 4743 4753
CC REPEAT 4754 4764
CC REPEAT 4765 4775
CC REPEAT 4776 4786
CC REPEAT 4787 4797
CC REPEAT 4798 4808
CC REPEAT 4809 4819
CC REPEAT 4820 4830
CC REPEAT 4831 4841
CC REPEAT 4842 4852
CC REPEAT 4853 4863
CC REPEAT 4864 4874
CC REPEAT 4875 4885
CC REPEAT 4886 4896
CC REPEAT 4897 4907
CC REPEAT 4908 4918
CC REPEAT 4919 4929
CC REPEAT 4930 4940
CC REPEAT 4941 4951
CC REPEAT 4952 4962
CC REPEAT 4963 4973
CC REPEAT 4974 4984
CC REPEAT 4985 4995
CC REPEAT 4996 5006
CC REPEAT 5007 5017
CC REPEAT 5018 5028
CC REPEAT 5029 5039
CC REPEAT 5040 5050
CC REPEAT 5051 5061
CC REPEAT 5062 5072
CC REPEAT 5073 5083
CC REPEAT 5084 5094
CC REPEAT 5095 5105
CC REPEAT 5106 5116
CC REPEAT 5117 5127
CC REPEAT 5128 5138
CC REPEAT 5139 5149
CC REPEAT 5150 5160
CC REPEAT 5161 5171
CC REPEAT 5172 5182
CC REPEAT 5183 5193
CC REPEAT 5194 5204
CC REPEAT 5205 5215
CC REPEAT 5216 5226
CC REPEAT 5227 5237
CC REPEAT 5238 5248
CC REPEAT 5249 5259
CC REPEAT 5260 5270
CC REPEAT 5271 5281
CC REPEAT 5282 5292
CC REPEAT 5293 5303
CC REPEAT 5304 5314
CC REPEAT 5315 5325
CC REPEAT 5326 5336
CC REPEAT 5337 5347
CC REPEAT 5348 5358
CC REPEAT 5359 5369
CC REPEAT 5370 5380
CC REPEAT 5381 5391
CC REPEAT 5392 5402
CC REPEAT 5403 5413
CC REPEAT 5414 5424
CC REPEAT 5425 5435
CC REPEAT 5436 5446
CC REPEAT 5447 5457
CC REPEAT 5458 5468
CC REPEAT 5469 5479
CC REPEAT 5480 5490
CC REPEAT 5491 5501
CC REPEAT 5502 5512
CC REPEAT 5513 5523
CC REPEAT 5524 5534
CC REPEAT 5535 5545
CC REPEAT 5546 5556
CC REPEAT 5557 5567
CC REPEAT 5568 5578
CC REPEAT 5579 5589
CC REPEAT 5590 5599
CC REPEAT 5600 5610
CC REPEAT 5611 5621
CC REPEAT 5622 5632
CC REPEAT 5633 5643
CC REPEAT 5644 5654
CC REPEAT 5655 5665
CC REPEAT 5666 5676
CC REPEAT 5677 5687
CC REPEAT 5688 5698
CC REPEAT 5699 5709
CC REPEAT 5710 5720
CC REPEAT 5721 5731
CC REPEAT 5732 5742
CC REPEAT 5743 5753
CC REPEAT 5754 5764
CC REPEAT 5765 5775
CC REPEAT 5776 5786
CC REPEAT 5787 5797
CC REPEAT 5798 5808
CC REPEAT 5809 5819
CC REPEAT 5820 5830
CC REPEAT 5831 5841
CC REPEAT 5842 5852
CC REPEAT 5853 5863
CC REPEAT 5864 5874
CC REPEAT 5875 5885
CC REPEAT 5886 5896
CC REPEAT 5897 5907
CC REPEAT 5908 5918
CC REPEAT 5919 5929
CC REPEAT 5930 5940
CC REPEAT 5941 5951
CC REPEAT 5952 5962
CC REPEAT 5963 5973
CC REPEAT 5974 5984
CC REPEAT 5985 5995
CC REPEAT 5996 6006
CC REPEAT 6007 6017
CC REPEAT 6018 6028
CC REPEAT 6029 6039
CC REPEAT 6040 6050
CC REPEAT 6051 6061
CC REPEAT 6062 6072
CC REPEAT 6073 6083
CC REPEAT 6084 6094
CC REPEAT 6095 6105
CC REPEAT 6106 6116
CC REPEAT 6117 6127
CC REPEAT 6128 6138
CC REPEAT 6139 6149
CC REPEAT 6150 6160
CC REPEAT 6161 6171
CC REPEAT 6172 6182
CC REPEAT 6183 6193
CC REPEAT 6194 6204
CC REPEAT 6205 6215
CC REPEAT 6216 6226
CC REPEAT 6227 6237
CC REPEAT 6238 6248
CC REPEAT 6249 6259
CC REPEAT 6260 6270
CC REPEAT 6271 6281
CC REPEAT 6282 6292
CC REPEAT 6293 6303
CC REPEAT 6304 6314
CC REPEAT 6315 6325
CC REPEAT 6326 6336
CC REPEAT 6337 6347
CC REPEAT 6348 6358
CC REPEAT 6359 6369
CC REPEAT 6370 6380
CC REPEAT 6381 6391
CC REPEAT 6392 6402
CC REPEAT 6403 6413
CC REPEAT 6414 6424
CC REPEAT 6425 6435
CC REPEAT 6436 6446
CC REPEAT 6447 6457
CC REPEAT 6458 6468
CC REPEAT 6469 6479
CC REPEAT 6480 6490
CC REPEAT 6491 6501
CC REPEAT 6502 6512
CC REPEAT 6513 6523
CC REPEAT 6524 6534
CC REPEAT 6535 6545
CC REPEAT 6546 6556
CC REPEAT 6557 6567
CC REPEAT 6568 6578
CC REPEAT 6579 6589
CC REPEAT 6590 6599
CC REPEAT 6600 6610
CC REPEAT 6611 6621
CC REPEAT 6622 6632
CC REPEAT 6633 6643
CC REPEAT 6644 6654
CC REPEAT 6655 6665
CC REPEAT 6666 6676
CC REPEAT 6677 6687
CC REPEAT 6688 6698
CC REPEAT 6699 6709
CC REPEAT 6710 6720
CC REPEAT 6721 6731
CC REPEAT 6732 6742
CC REPEAT 6743 6753
CC REPEAT 6754 6764
CC REPEAT 6765 6775
CC REPEAT 6776 6786
CC REPEAT 6787 6797
CC REPEAT 6798 6808
CC REPEAT 6809 6819
CC REPEAT 6820 6830
CC REPEAT 6831 6841
CC REPEAT 6842 6852
CC REPEAT 6853 6863
CC REPEAT 6864 6874
CC REPEAT 6875 6885
CC REPEAT 6886 6896
CC REPEAT 6897 6907
CC REPEAT 6908 6918
CC REPEAT 6919 6929
CC REPEAT 6930 6940
CC REPEAT 6941 6951
CC REPEAT 6952 6962
CC REPEAT 6963 6973
CC REPEAT 6974 6984
CC REPEAT 6985 6995
CC REPEAT 6996 7006
CC REPEAT 7007 7017
CC REPEAT 7018 7028
CC REPEAT 7029 7039
CC REPEAT 7040 7050
CC REPEAT 7051 7061
CC REPEAT 7062 7072
CC REPEAT 7073 7083
CC REPEAT 7084 7094
CC REPEAT 7095 7105
CC REPEAT 7106 7116
CC REPEAT 7117 7127
CC REPEAT 7128 7138
CC REPEAT 7139 7149
CC REPEAT 7150 7160
CC REPEAT 7161 7171
CC REPEAT 7172 7182
CC REPEAT 7183 7193
CC REPEAT 7194 7204
CC REPEAT 7205 7215
CC REPEAT 7216 7226
CC REPEAT 7227 7237
CC REPEAT 7238 7248
CC REPEAT 7249 7259
CC REPEAT 7260 7270
CC REPEAT 7271 7281
CC REPEAT 7282 7292
CC REPEAT 7293 7303
CC REPEAT 7304 7314
CC REPEAT 7315 7325
CC REPEAT 7326 7336
CC REPEAT 7337 7347
CC REPEAT 7348 7358
CC REPEAT 7359 7369
CC REPEAT 7370 7380
CC REPEAT 7381 7391
CC REPEAT 7392 7402
CC REPEAT 7403 7413
CC REPEAT 7414 7424
CC REPEAT 7425 7435
CC REPEAT 7436 7446
CC REPEAT 7447 7457
CC REPEAT 7458 7468
CC REPEAT 7469 7479
CC REPEAT 7480 7490
CC REPEAT 7491 7501
CC REPEAT 7502 7512
CC REPEAT 7513 7523
CC REPEAT 7524 7534
CC REPEAT 7535 7545
CC REPEAT 7546 7556
CC REPEAT 7557 7567
CC REPEAT 7568 7578
CC REPEAT 7579 7589
CC REPEAT 7590 7599
CC REPEAT 7600 7610
CC REPEAT 7611 7621
CC REPEAT 7622 7632
CC REPEAT 7633 7643
CC REPEAT 7644 7654
CC REPEAT 7655 7665
CC REPEAT 7666 7676
CC REPEAT 7677 7687
CC REPEAT 7688 7698
CC REPEAT 7699 7709
CC REPEAT 7710 7720
CC REPEAT 7721 7731
CC REPEAT 7732 7742
CC REPEAT 7743 7753
CC REPEAT 7754 7764
CC REPEAT 7765 7775
CC REPEAT 7776 7786
CC REPEAT 7787 7797
CC REPEAT 7798 7808
CC REPEAT 7809 7819
CC REPEAT 7820 7830
CC REPEAT 7831 7841
CC REPEAT 7842 7852
CC REPEAT 7853 7863
CC REPEAT 7864 7874
CC REPEAT 7875 7885
CC REPEAT 7886 7896
CC REPEAT 7897 7907
CC REPEAT 7908 7918
CC REPEAT 7919 7929
CC REPEAT 7930 7940
CC REPEAT 7941 7951
CC REPEAT 7952 7962
CC REPEAT 7963 7973
CC REPEAT 7974 7984
CC REPEAT 7985 7995
CC REPEAT 7996 8006
CC REPEAT 8007 8017
CC REPEAT 8018 8028
CC REPEAT 8029 8039
CC REPEAT 8040 8050
CC REPEAT 8051 8061
CC REPEAT 8062 8072
CC REPEAT 8073 8083
CC REPEAT 8084 8094
CC REPEAT 8095 8105
CC REPEAT 8106 8116
CC REPEAT 8117 8127
CC REPEAT 8128 8138
CC REPEAT 8139 8149
CC REPEAT 8150 8160
CC REPEAT 8161 8171
CC REPEAT 8172 8182
CC REPEAT 8183 8193
CC REPEAT 8194 8204
CC REPEAT 8205 8215
CC REPEAT 8216 8226
CC REPEAT 8227 8237
CC REPEAT 8238 8248
CC REPEAT 8249 8259
CC REPEAT 8260 8270
CC REPEAT 8271 8281
CC REPEAT 8282 8292
CC REPEAT 8293 8303
CC REPEAT 8304 8314
CC REPEAT 8315 8325
CC REPEAT 8326 8336
CC REPEAT 8337 8347
CC REPEAT 8348 8358
CC REPEAT 8359 8369
CC REPEAT 8370 8380
CC REPEAT 8381 8391
CC REPEAT 8392 8402
CC REPEAT 8403 8413
CC REPEAT 8414 8424
CC REPEAT 8425 8435
CC REPEAT 8436 8446
CC REPEAT 8447 8457
CC REPEAT 8458 8468
CC REPEAT 8469 8479
CC REPEAT 8480 8490
CC REPEAT 8491 8501
CC REPEAT 8502 8512
CC REPEAT 8513 8523
CC REPEAT 8524 8534
CC REPEAT 8535 8545
CC REPEAT 8546 8556
CC REPEAT 8557 8567
CC REPEAT 8568 8578
CC REPEAT 8579 8589
CC REPEAT 8590 8599
CC REPEAT 8600 8610
CC REPEAT 8611 8621
CC REPEAT 8622 8632
CC REPEAT 8633 8643
CC REPEAT 8644 8654
CC REPEAT 8655 8665
CC REPEAT 8666 8676
CC REPEAT 8677 8687
CC REPEAT 8688 8698
CC REPEAT 8699 8709
CC REPEAT 8710 8720
CC REPEAT 8721 8731
CC REPEAT 8732 8742
CC REPEAT 8743 8753
CC REPEAT 8754 8764
CC REPEAT 8765 8775
CC REPEAT 8776 8786
CC REPEAT 8787 8797
CC REPEAT 8798 8808
CC REPEAT 8809 8819
CC REPEAT 8820 8830
CC REPEAT 8831 8841
CC REPEAT 8842 8852
CC REPEAT 8853 8863
CC REPEAT 8864 8874
CC REPEAT 8875 8885
CC REPEAT 8886 8896
CC REPEAT 8897 8907
CC REPEAT 8908 8918
CC REPEAT 8919 8929
CC REPEAT 8930 8940
CC REPEAT 8941 8951
CC REPEAT 8952 8962
CC REPEAT 8963 8973
CC REPEAT 8974 8984
CC REPEAT 8985 8995
CC REPEAT 8996 9006
CC REPEAT 9007 9017
CC REPEAT 9018 9028
CC REPEAT 9029 9039
CC REPEAT 9040 9050
CC REPEAT 9051 9061
CC REPEAT 9062 9072
CC REPEAT 9073 9083
CC REPEAT 9084 9094
CC REPEAT 9095 9105
CC REPEAT 9106 9116
CC REPEAT 9117 9127
CC REPEAT 9128 9138
CC REPEAT 9139 9149
CC REPEAT 9150 9160
CC REPEAT 9161 9171
CC REPEAT 9172 9182
CC REPEAT 9183 9193
CC REPEAT 9194 9204
CC REPEAT 9205 9215
CC REPEAT 9216 9226
CC REPEAT 9227 9237
CC REPEAT 9238 9248
CC REPEAT 9249 9259
CC REPEAT 9260 9270
CC REPEAT 9271 9281
CC REPEAT 9282 9292
CC REPEAT 9293 9303
CC REPEAT 9304 9314
CC REPEAT 9315 9325
CC REPEAT 9326 9336
CC REPEAT 9337 9347
CC REPEAT 9348 9358
CC REPEAT 9359 9369
CC REPEAT 9370 9380
CC REPEAT 9381 9391
CC REPEAT 9392 9402
CC REPEAT 9403 9413
CC REPEAT 9414 9424
CC REPEAT 9425 9435
CC REPEAT 9436 9446
CC REPEAT 9447 9457
CC REPEAT 9458 9468
CC RE

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 11:05:23 ; Search time 128 Seconds
(without alignments)
1604.706 Million cell updates/sec

Title: US-10-019-495-9
Perfect score: 3473
Sequence: 1 MKNVPVPVCRPLVEKDPMT.....VLSKAERSHIIVQSVTYPE 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3191	91.9	1328	11 Q8CHB5	Q8chb5 mus musculus
2	3057	88.0	599	6 Q95LM5	Q95lm5 macaca fasc
3	2870	82.6	587	11 Q8CAB6	Q8cab6 mus musculus
4	1892	54.5	827	11 Q8C7W0	Q8c7w0 mus musculus
5	1821.5	52.4	1307	11 Q8CJC2	Q8cjc2 mus musculus
6	1814.5	52.2	1179	4 Q60271	Q60271 homo sapien
7	1814.5	52.2	1311	4 Q8IZK7	Q8izk7 homo sapien
8	943	27.2	734	4 Q9H811	Q9h811 homo sapien
9	623.5	18.0	200	4 Q7Z508	Q7z508 homo sapien
10	460.5	13.3	766	4 Q60905	Q60905 homo sapien
11	460.5	13.3	780	4 Q86WC8	Q86wc8 homo sapien
12	449	12.9	1082	11 Q80U35	Q80u35 mus musculus
13	442.5	12.7	1510	4 Q9Y4G3	Q9y4g3 homo sapien
14	442.5	12.7	2063	4 Q96PE2	Q96pe2 homo sapien
15	436.5	12.6	1101	5 Q9VEF9	Q9vef9 drosophila
16	389.5	11.2	760	6 Q7YRB3	Q7yrb3 papio hamad

17	333.5	9.6	1284	4 Q9HCE6	Q9hce6 homo sapien
18	320	9.2	720	11 Q8VDH5	Q8vdh5 mus musculus
19	320	9.2	1241	11 Q8CIA1	Q8cia1 mus musculus
20	306	8.8	988	4 Q14665	Q14665 homo sapien
21	305.5	8.8	1405	4 Q15013	Q15013 homo sapien
22	291.5	8.4	701	11 Q80VH8	Q80vh8 mus musculus
23	291.5	8.4	1139	11 Q8C033	Q8c033 mus musculus
24	265	7.6	1458	5 Q9VX48	Q9vx48 drosophila
25	265	7.6	1740	5 Q86B45	Q86b45 drosophila
26	254.5	7.3	755	5 Q8IGL6	Q8igl6 drosophila
27	233	6.7	596	4 Q9NVT3	Q9nvt3 homo sapien
28	210	6.0	193	11 Q8BZV4	Q8bzv4 mus musculus
29	146.5	4.2	249	16 Q9RDN7	Q9rdn7 streptomyce
30	146.5	4.2	985	12 Q67643	Q67643 gallid herp
31	145	4.2	483	12 Q83344	Q83344 murid herpe
32	144.5	4.2	1054	11 Q9JM99	Q9jtm99 mus musculus
33	144.5	4.2	1527	3 Q8WZL5	Q8wzl5 yarrowia li
34	144.5	4.2	2344	5 Q9N3Y8	Q9n3y8 caenorhabdi
35	142	4.1	300	10 Q07373	Q07373 chlamydomon
36	141.5	4.1	582	16 Q82F59	Q82f59 streptomyce
37	140	4.0	483	12 Q9DYE3	Q9dye3 murine herp
38	140	4.0	647	10 Q8S148	Q8s148 oryza sativ
39	138.5	4.0	464	16 Q8FS16	Q8fs16 corynebacte
40	138.5	4.0	667	16 Q9RJY5	Q9rjy5 streptomyce
41	138.5	4.0	825	5 Q7YZW4	Q7yzw4 caenorhabdi
42	138.5	4.0	1032	5 P91365	P91365 caenorhabdi
43	138.5	4.0	2232	5 Q8IFX6	Q8ifx6 caenorhabdi
44	138	4.0	586	16 Q98J75	Q98j75 thizobium l
45	138	4.0	1013	16 Q9KE85	Q9ke85 bacillus ha

ALIGNMENTS

RESULT 1

Q8CHB5 ID Q8CHB5 PRELIMINARY; PRT; 1328 AA.
AC Q8CHB5, 2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 24, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 25, Last annotation update)
DE MKIAA1066 protein (Fragment)
GN MAPK8IP3 OR MKIAA1066.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Hara Y., Nagase T.,
RA Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT I. The complete nucleotide sequences of 100 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries."
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB093282; BAC41466.2;
DR MGD; MGI:1353598; Mapk8ip3.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005078; F:MAP-kinase scaffold activity; IDA.
DR GO; GO:0007254; P:JNK cascade; IDA.
FT NON_TER 1
SQ SEQUENCE 1328 AA; 146328 MW; 647EF78FC762B29F CRC64;

Query Match 91.9%; Score 3191; DB 11; Length 1328;
Best Local Similarity 91.2%; Pred. No. 7e-223;
Matches 594; Conservative 24; Mismatches 33; Indels 0; Gaps 0;
Qy 1 MKNVPVPVCRPLVEKDPMTKLWCAAGVNLSGWRPNEDDAGNKGKAPGRDPLTCDREGD 60
Db 678 MKNVPVPVCRPLVEKDPSTKLWCAAGVNLSGWKPHEDSSNGKPVFGRDPLTCDREGE 737

```

QY 61 GEPKSAHTSPEKKAKELPEMDATSSRWILSTLTTSKVVIIIDANOGTVVDQFTVCNA 120
DB 738 GEPKSAHTSPEKKAKELPEMDATSSRWILSTLTTSKVVIIIDANOGTVVDQFTVCNA 797
QY 121 HVLICISSIPAASDSYPPGEMFLSDVNPDPGADGVLGAGITLVGCATRCNVPRSNCSRR 180
DB 798 HVLICISSIPAASDSYPPGEMFLSDVNPDPGADGVLGAGITLVGCATRCNVPRSNCSRR 857
QY 181 GDTPLDKGQGEVATIANGVNPSQSTEEATEATEVDPGSPBETATLRPGPLTEHVFT 240
DB 858 GDTPLDKGQGEVATIANGVNPSQSTEEATEATEVDPGSPBETATLRPGPLTEHVFT 917
QY 241 DPAPTPSGPOPGSENGEPDSSSTRPEPSGPTGAGSSAAPTMLGAQGMWLYVHSA 300
DB 918 DPAPTPSGPOPGSENGEPDSSSTRPEPSGPTGAGSSAAPTMLGAQGMWLYVHSA 977
QY 301 VANWKKCLHSIKLSDSVLSLVHVGRVLVALADGTLAIFHRGEGQWDLNHYHMLDLGHP 360
DB 978 VANWKKCLHSIKLSDSVLSLVHVGRVLVALADGTLAIFHRGEGQWDLNHYHMLDLGHP 1037
QY 361 HHSIRCMVAVVDRVWCVGNKXVHVIQPKTMOIEKSFDAHPRESQVROLAWIGDVWVSI 420
DB 1038 HHSIRCMVAVVDRVWCVGNKXVHVIQPKTMOIEKSFDAHPRESQVROLAWIGDVWVSI 1097
QY 421 RLDSTLRLYHAHTHOHLQDQVDIEPVVSKMLGTGKLGFSFVRITALLVAGSLWVGTNGV 480
DB 1098 RLDSTLRLYHAHTHOHLQDQVDIEPVVSKMLGTGKLGFSFVRITALLVAGSLWVGTNGV 1157
QY 481 VISIPLTETVVLHRLGQLGLRANKTSPSTSGEGARPGGIIHVYGDSSDRAASSFTPYCSM 540
DB 1158 VISIPLTETVVLHRLGQLGLRANKTSPSTSGEGARPGGIIHVYGDSSDRAASSFTPYCSM 1217
QY 541 AQALCFHGHRDAVKFFVSPGNVLATLNGSVLSDPAEGPAPASVEGQKLNVLVL 600
DB 1218 AQALCFHGHRDAVKFFVSPGNVLATLNGSVLSDPAEGPAPASVEGQKLNVLVL 1277
QY 601 SGGEGYIDFRIGDGRDDETERGAGDMQVSKPVLKASERSHIIVQVSYTPE 651
DB 1278 SGGEGYIDFRIGDGRDDETERGAGDMQVSKPVLKASERSHIIVQVSYTPE 1328

RESULT 2
Q95LW5 PRELIMINARY; PRT; 599 AA.
AC Q95LW5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]_TaxID=9541;
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA Terao K., Sugano S.;
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
RT libraries."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071076; BAB64469.1; -.
KW Hypothetical protein.
SQ SEQUENCE 599 AA; 64132 MW; 565063493B8EC193 CRC64;
Query Match 88.0%; Score 3057; DB 6; Length 599;
Best Local Similarity 98.3%; Pred. No. 1.3e-213;
Matches 571; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 61 GEPKSAHTSPEKKAKELPEMDATSSRWILSTLTTSKVVIIIDANOGTVVDQFTVCNA 120
DB 61 GEPKSAHTSPEKKAKELPEMDATSSRWILSTLTTSKVVIIIDANOGTVVDQFTVCNA 120
QY 121 HVLICISSIPAASDSYPPGEMFLSDVNPDPGADGVLGAGITLVGCATRCNVPRSNCSRR 180
DB 121 HVLICISSIPAASDSYPPGEMFLSDVNPDPGADGVLGAGITLVGCATRCNVPRSNCSRR 180
QY 181 GDTPLDKGQGEVATIANGVNPSQSTEEATEATEVDPGSPBETATLRPGPLTEHVFT 240
DB 181 GDTPLDKGQGEVATIANGVNPSQSTEEATEATEVDPGSPBETATLRPGPLTEHVFT 240
QY 241 DPAPTPSGPOPGSENGEPDSSSTRPEPSGPTGAGSSAAPTMLGAQGMWLYVHSA 300
DB 241 DPAPTPSGPOPGSENGEPDSSSTRPEPSGPTGAGSSAAPTMLGAQGMWLYVHSA 300
QY 301 VANWKKCLHSIKLSDSVLSLVHVGRVLVALADGTLAIFHRGEGQWDLNHYHMLDLGHP 360
DB 301 VANWKKCLHSIKLSDSVLSLVHVGRVLVALADGTLAIFHRGEGQWDLNHYHMLDLGHP 360
QY 361 HHSIRCMVAVVDRVWCVGNKXVHVIQPKTMOIEKSFDAHPRESQVROLAWIGDVWVSI 420
DB 361 HHSIRCMVAVVDRVWCVGNKXVHVIQPKTMOIEKSFDAHPRESQVROLAWIGDVWVSI 420
QY 421 RLDSTLRLYHAHTHOHLQDQVDIEPVVSKMLGTGKLGFSFVRITALLVAGSLWVGTNGV 480
DB 421 RLDSTLRLYHAHTHOHLQDQVDIEPVVSKMLGTGKLGFSFVRITALLVAGSLWVGTNGV 480
QY 481 VISIPLTETVVLHRLGQLGLRANKTSPSTSGEGARPGGIIHVYGDSSDRAASSFTPYCSM 540
DB 481 VISIPLTETVVLHRLGQLGLRANKTSPSTSGEGARPGGIIHVYGDSSDRAASSFTPYCSM 540
QY 541 AQALCFHGHRDAVKFFVSPGNVLATLNGSVLSDPAEGP 581
DB 541 AQALCFHGHRDAVKFFVSPGNVLATLNGSVLSDPAEPLG 581

RESULT 3
Q8CAB6 PRELIMINARY; PRT; 587 AA.
AC Q8CAB6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mitogen-activated protein kinase 8 interacting protein 3
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK039130; BAC30247.1; -.
FT NON_TER 1
SQ SEQUENCE 587 AA; 62680 MW; BCA08B87D724CBD7 CRC64;
Query Match 82.6%; Score 2870; DB 11; Length 587;
Best Local Similarity 92.0%; Pred. No. 4.9e-200;
Matches 540; Conservative 18; Mismatches 29; Indels 0; Gaps 0;

```

```

QY 65 SAHTSPEKKAKELPEMDATSSRWILSTLTTSKVVIIIDANOGTVVDQFTVCNAHVL 124
DB 1 STHPSPEKKAKETEADATSSRWILSTLTTSKVVIIIDANOGTVVDQFTVCNAHVL 60
QY 125 ISSIPAASDSYPPGEMFLSDVNPDPGADGVLGAGITLVGCATRCNVPRSNCSRGDTP 184

```

```

Db 61 ISSIPAAASDSDYPPGEMFLDSVNDPDSGADGLAGITLVGCATRCNVPKSCSSRGDTP 120
QY 185 VLDKGGQGVATANGKVNPSQSTTEATEVDPDPGPPEPATATLPGLTEHVETDPA 244
Db 121 VLDKGGQGVATANGKVNPSQSTTEATEVDPDPGPPEPATATLPGLTEHVETDPA 180
QY 245 TPSSGPQSGSENGPEPSSSTRPEPEPDDPTGAGSSAAPTMMWLGAQNGWLYVHSANW 304
Db 181 TPSSSTQASENGSENGTIVQVPEPSEGLSTTSSAAPTMMWLGAQNGWLYVHSANW 240
QY 305 KKCLHSIKLDSVLSLVHVKGRVLVALADGTALAIHFRGEGDQWDLNHYHMLDLGHPHHSI 364
Db 241 KKCLHSIKLDSVLSLVHVKGRVLVALADGTALAIHFRGEGDQWDLNHYHMLDLGHPHHSI 300
QY 365 RCMAVYDVRVWCGYKYNKHVHVPKTMQIEKSFDAHPRESQVRLAWTGDGVVWSIRLDS 424
Db 301 RCMAVYDVRVWCGYKYNKHVHVPKTMQIEKSFDAHPRESQVRLAWTGDGVVWSIRLDS 360
QY 425 TRLRYHAHTHQLQDVLDIEPYVSKMLGTGKLGFSFVRITALLVAGSRLWVGTVGNGVLSI 484
Db 361 TRLRYHAHTHQLQDVLDIEPYVSKMLGTGKLGFSFVRITALLVAGSRLWVGTVGNGVLSI 420
QY 485 PLTEVTVLHRCQLLGLRANKTSPSGEGARPGGIIHVYGGDSDRAASSFIPYCSMAQAQ 544
Db 421 PLTEVTVLHRCQLLGLRANKTSPSGEGARPGGIIHVYGGDSDRAASSFIPYCSMAQAQ 480
QY 545 LCFPHGRDAVFFVSPGNVATLNGSVLDSPAGPGPAAPASEVGGKLRNLVLSGGE 604
Db 481 LCFPHGRDAVFFVSPGNVATLNGSVLDSPAGPGPAAPASEVGGKLRNLVLSGGE 540
QY 605 GYIDFRIGDGEDDETEGAGDMSQVKKPVLKASERSHIIWQVSTYPE 651
Db 541 GYIDFRIGDGEDDETEGAGDMSQVKKPVLKASERSHIIWQVSTYPE 587

RESULT 4
Q8C7W0 PRELIMINARY; PRT; 827 AA.
AC Q8C7W0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Similar to sperm surface protein.
GN SPAG9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12456851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK049161; BAC33577.1; -.
DR MGD; MGI:1918084; Spag9.
SQ SEQUENCE 827 AA; 90652 MW; 06627157DDC785A0 CRC64;

Query Match 54.5%; Score 1892; DB 11; Length 827;
Best Local Similarity 57.8%; Pred. No. 8.7e-129;
Matches 391; Conservative 96; Mismatches 147; Indels 42; Gaps 17;

QY 1 MNKVPVPCRPVLEKDPKMLCAAGVNLGWRPNEDDAGNVKAP-GRDPLTCDREG 59
Db 159 MNKLPVPLRPLDEKDKAMKLCVAVNLGSGKTR--DGGSVVGASVFKYKTAGLDTG 216

QY 60 DGPFSKNTSPE-----KKKAKELPMDATSGRWLTSTLTSTSKVVIIDANQPGTWD 113
Db 217 SKORSASQSLDKLDQLKEQKEFKNQBELSSQVWICTSTHTKVIIDAVQPGNILD 276

```

```

QY 114 QFTVCNAHVLCISSIPAAASDSDYPPGEMFLDS-----VNPDPGADGVLAGIT 162
Db 277 SFTVCNSHVLCTASVPGARETDYPAGEBELSBSQVDKASLCGSMTSNSAEMDSLGGIT 336
QY 163 LVGCATR--CNVPRNSCSRGTTPVLDKGGQGVATANGKVNPSQST--EATEATEVDPGP 220
Db 337 VVGSTEGTLTAATSPSTNGASPVIEK--PPMET--ENSEVDENIPTAEATEATE--GNAG 393
QY 221 PSEPETATLRGPPLTEHVETDPAFTP-----SSGPQSGSENGPEPSSSTRP--EPEPSD 274
Db 394 STEDTVDISQGVVTEHVETDPLGVQIPEDLSPVFQSSNDSVDYKQISVLPNEQDLARE 453
QY 275 PTGAGSNAAPTMMWLGAQNGWLYVHSANWKKLHSLIKLDSVLSLVHVKGRVLVALADG 334
Db 454 EAQKXSSLLPTMMWLGAQNGCLYVHSSVAQWRKCLHSLIKLDSILSVHVKGIVLVALADG 513
QY 335 TLAIFHRGEGDQWDLNHYHMLDLGHPHHSIRCMVAVYDVRVWCGYKYNKHVHVPKTMQIEK 394
Db 514 TLAIFHRGEGDQWDLNHYHMLDLGHPHHSIRCMVAVYDVRVWCGYKYNKHVHVPKTMQIEK 394
QY 395 SFDAPRRRESQVRLAWTGDGVVWSIRLDSRLRYHAHTHQLQDVLDIEPYVSKMLGTGK 454
Db 574 SFDAPRRRESQVRLAWTGDGVVWSIRLDSRLRYHAHTHQLQDVLDIEPYVSKMLGTGK 454
QY 455 LGFSFVRITALLVAGSRLWVGTVGNGVLSIPLTEVTVLHRCQLLGLRANKTSPSGEGAR 514
Db 634 LGFSFVRITALLVAGSRLWVGTVGNGVLSIPLTEVTVLHRCQLLGLRANKTSPSGEGAR 514
QY 515 PGGIIHVYGGDSDRAA--SSFIPYCSMAQAQLCFPHGRDAVFFVSPGNVLSI-----ATLNG 570
Db 692 PGSVIRVYGGDSDRAA--SSFIPYCSMAQAQLCFPHGRDAVFFVSPGNVLSI-----ATLNG 570
QY 571 SVLDSPAEGPGPAAPASEVGGKLRNLVLSGSGGYIDFRIGDGEDDETEGAGDMSQV 630
Db 752 AGLTADKAGSSAQEPSSQT---PLKSLMVLISGSGGYIDFRIGD--EGGESEL--LGEDUPL 806
QY 631 PVLKASERSHIIWQV 646
Db 807 PSVTKASERSHIIWQV 822

RESULT 5
Q8CJC2 PRELIMINARY; PRT; 1307 AA.
AC Q8CJC2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE JNK-associated leucine-zipper protein.
GN SPAG9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RX MEDLINE=22295031; PubMed=12391307;
RA Lee C.M., Onesime D., Reddy C.D., Dhanasekaran N., Reddy E.P.;
RT "JLP: A scaffolding protein that tethers JNK/p38MAPK signaling modules
RT and transcription factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14189-14194(2002).
DR EMBL; AF327451; AAN61564.1; -.
DR MGD; MGI:1918084; Spag9.
SQ SEQUENCE 1307 AA; 144592 MW; E56C26AA9DD8278C CRC64;

Query Match 52.4%; Score 1821.5; DB 11; Length 1307;
Best Local Similarity 56.4%; Pred. No. 2.1e-123;
Matches 381; Conservative 93; Mismatches 147; Indels 55; Gaps 18;

QY 1 MNKVPVPCRPVLEKDPKMLCAAGVNLGWRPNEDDAGNVKAP-GRDPLTCDREG 59
Db 652 MNKLPVPLRPLDEKDKAMKLCVAVNLGSGKTR--DGGSVVGASVFKYKTAGLDTG 709

```


[illegible]

```
Matches 128; Conservative 31; Mismatches 37; Indels 11; Gaps 6;
QY 449 MLGTGKLGFSVRITALLVAGSRVWGTGNGVTSIPLTETVVLHRRGOLLGLRANKTSPT 508
D 1 MLGTGKLGFSVRITALLVAGSRVWGTGNGVTSIPLTETVVLHRRGOLLGLRANKTSV 60
QY 509 SGEARGGIHHVYGDSSDRAA--SSFTPTCSMAQAOLCFHGRDVAVKFFVSPGNVL-- 565
D 61 PGN--RFGSVIRVYGDSSDKVTGTFIPYCSMAHAQLCFHGRDVAVKFFVAVPGQVISP 118
QY 566 -ATLNGSVLDPAGPGFAAPASAVEGOKLRNVLVLSGEGYIDFRIGDGEDDETEGAG 624
D 119 QSSSGTDLTGKAGPSAQEGSQT---PLAKMLVSGEGYIDFRMCD-EGGSEL-LG 173
QY 625 DMSQKVPVLSKAERSHIIVQVSTYTP 651
D 174 EDLPLEPVSVTAKERSHLIVQVMYGYNE 200
RESULT 10
O60905 PRELIMINARY; PRT; 766 AA.
ID AC O60905;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sperm specific protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Shankar S., Mohapatra B., Suri A.;
RT "Cloning of a novel human testis mRNA specifically expressed in
RT testicular haploid germ cells, having unique palindromic sequences and
RT encoding a leucine zipper dimerization motif.,"
RL Biochem. Biophys. Res. Commun. 243:561-565(1998).
DR EMBL; X91879; CAA62987.1; -.
DR PIR; JC5958;
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0007283; P:spermatogenesis; TAS.
DR GO; GO:0007283; P:spermatogenesis; TAS.
SQ SEQUENCE 766 AA; 83858 MW; C09E76A583604A68 CRC64;
```

```
Query Match 13.3%; Score 460.5; DB 4; Length 766;
Best Local Similarity 42.7%; Pred. No. 7.8e-25;
Matches 112; Conservative 40; Mismatches 85; Indels 25; Gaps 8;
QY 1 MKNVPVPCRPVLEKDTMKLWCAAGVNLGWRPNEDDAGNGVKPAP-GRDPLTCDREG 59
D 487 MKNLPVPCRPVLEKDTMKLWCAAGVNLGWRPNEDDAGNGVKPAP-GRDPLTCDREG 544
QY 60 DGEPKSAHTSP-----KKKAKELPEMDATSSRVWLTSTLTTSKVLIIDANQPGTVVD 113
D 545 SKQRSASQSLDKLDQELKEQKELNQELSSLVWICTSHSATKVLIIIDAVQGNIL 604
QY 114 QFTVCNAHVLCISSIPAAASDSYPPCEMFLDSD-----VNPEDPGADGVLGIT 162
D 605 SFTVCNHLVLCIASVPGARETDYPAGEDLSSEGQVDKASLCGSMTSNSAETDSLGGIT 664
QY 163 LVGC-ATRCNVPNSCSRGDTPVLIDKGGQEVATTANGKVNPSQST-BEATEATEVPDPG 220
D 665 VVGSAGVGTGATSPSTNGASPVMDKPPMEA--ENSEVDENVPTAEATEATE-GNAG 721
QY 221 PSEPETATLRPGPLTEHVFTDP 242
D 722 SAEDTVDISQTGVYTHEVFTDP 743
```

```
RESULT 11
Q86WC8
```

```
ID O86WC8 PRELIMINARY; PRT; 780 AA.
AC O86WC8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PHE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Yasuoka H., Ihn H., Medsger T.A. Jr., Hirakata M., Kawakami Y.,
RA Ikeda Y., Kuwana M.;
RT "A novel testicular antigen targeted by autoantibodies in systemic
RT scleriosis: upregulated expression in affected dermal fibroblasts.,"
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY219897; AAO66462.1; -.
DR SEQUENCE 780 AA; 85309 MW; ABA29417B34A8875 CRC64;
Query Match 13.3%; Score 460.5; DB 4; Length 780;
Best Local Similarity 42.7%; Pred. No. 8e-25;
Matches 112; Conservative 40; Mismatches 85; Indels 25; Gaps 8;
QY 1 MKNVPVPCRPVLEKDTMKLWCAAGVNLGWRPNEDDAGNGVKPAP-GRDPLTCDREG 59
D 501 MKNLPVPCRPVLEKDTMKLWCAAGVNLGWRPNEDDAGNGVKPAP-GRDPLTCDREG 558
QY 60 DGEPKSAHTSP-----KKKAKELPEMDATSSRVWLTSTLTTSKVLIIDANQPGTVVD 113
D 559 SKQRSASQSLDKLDQELKEQKELNQELSSLVWICTSHSATKVLIIIDAVQGNIL 618
QY 114 QFTVCNAHVLCISSIPAAASDSYPPCEMFLDSD-----VNPEDPGADGVLGIT 162
D 619 SFTVCNHLVLCIASVPGARETDYPAGEDLSSEGQVDKASLCGSMTSNSAETDSLGGIT 678
QY 163 LVGC-ATRCNVPNSCSRGDTPVLIDKGGQEVATTANGKVNPSQST-BEATEATEVPDPG 220
D 679 VVGSAGVGTGATSPSTNGASPVMDKPPMEA--ENSEVDENVPTAEATEATE-GNAG 735
QY 221 PSEPETATLRPGPLTEHVFTDP 242
D 736 SAEDTVDISQTGVYTHEVFTDP 757
RESULT 12
Q80U35 PRELIMINARY; PRT; 1082 AA.
AC Q80U35;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MKIAA0337 protein (Fragment).
GN MKIAA0337.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Kosa H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.,"
RL DNA Res. 10:35-48(2003).
DR EMBL; AK122250; BAC65532.1; -.
DR InterPro; IPR000219; RhoGEF.
DR Pfam; PF00621; RhoGEF; 1.
```


GN TEM4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21443268; PubMed=11559528; Nanda A., Vogelstein B.,
RA Carson-Walter E.B., Watkins D.N., Nanda A., Vogelstein B.,
RT Kinzler K.W., St Croix B.
RT "Cell surface tumor endothelial markers are conserved in mice and
RT humans";
RL Cancer Res. 61:6649-6655 (2001).
DR EMBL; AF378754; AAL11991.1; -.
DR InterPro; IPR000219; RhoGEF.
DR Pfam; PF00621; RhoGEF; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS50010; DH 2; 1.
SQ SEQUENCE 2063 AA; 221671 MW; 76A53FP6CBF5F5C0 CRC64;

Query Match 12.7%; Score 442.5; DB 4; Length 2063;
Best Local Similarity 26.4%; Pred. No. 6e-23;
Matches 168; Conservative 86; Mismatches 223; Indels 159; Gaps 25;

Qy 85 SSRWILSTLTTSKVLIIDANQGTGVVDTVCNAHVLCISSIPAAASDSY--PPGEMF 142
Db 1513 SPRVWNCSDGYGVQVCIMSLHPEFNTSCNGVNRILCVASPAYSSA----- 1572
Qy 143 LDSVNPDPGADGVLGITVGCATRCNVRPNSC-----SSRGDTFPVLDDK 189
Db 1573 SPPTAPFPAGELDVEAADEAATLAEPGPCLHISIAGSGLEMTFGLGEGDPREL 1632
Qy 190 -----QGEVATIANGVN---PS--QSTEEATEATEVDPDPGSE 223
Db 1633 VPFDSDSDSSPSPSGTLQASRSTISSFGNEPTSSKEATAETTSSEEQEPG-FL 1691
Qy 224 PETATLRPGPLTEHVFTDPAPTSSGGPQPGSENGPEPPDSSTRPEP--SGDPTGAGSS 281
Db 1692 PLSGSFGPG-----GPCGT---SPMDGALRRSHSGFTRGSLDLSVDPEAYQSS 1740
Qy 282 AAPTMTLGAQNGWLVHSAVANWKKLHSIKKD--SVLSLVHVKRVLVALADGTLAIF 339
Db 1741 ----VVLGTEDGCVHVYQSSDIRDRNSMKLQHAASVTCILYLNQNVFVSLANGELVY 1796
Qy 340 HRGEGQNDLNYHLMIDLGHPHHSITRCMAVVDVRCVGVKQKHVVIQKTIQIEKSPDAH 399
Db 1797 QREAGHFDPQNFKSVTLGTQSPITKVMVSVGGRLWCQGNRVLVSPDTTLQLEHMFYV- 1855
Qy 400 PRRESQVQLAWIGD--GVVSVIRLSTLRLYHAHTHQLQDVVDIEPVVSKML-GTGKL 455
Db 1856 --GQSSRCVACWVDSGLGVVTLKGSARVCLYHPDTFEQLAEVDVTPVHRMLAGSDAI 1913
Qy 456 ---GFSEFVRITALLVAGSRMLVWGTVNGVVISIPLTE-TVVLHRLQLGLRANKTSPGSG 510
Db 1914 LRQKAAACLRITALLVCEBLLWVGTSGVLTMTSPGTVCPRAPL-----SPT-G 1964
Qy 511 EGAREGGIHHVYGDSSDRASSFPYCSMAQAQLCFHGHRDAVKFFVSV----- 560
Db 1965 LG-----QGTGTHVRFLAAVQLVQSGEGYIDFR 1989
Qy 561 -----PGNV---LATNGSVLDSP-ABGPGPAAPASEVEGQKRNVLVLSGGEGYIDFR 610
Db 1990 PTPPPPTGPKLPSLEHR--DSPWHRGPAAPAP-----KMLVISGGDGYEDFR 2037
Qy 611 IGDGDEDETEGAGDMSQVKPVLKSAERSHIIVQV 646
Db 2038 LSSGGSSSETVGRDSS-----TNHLLNRV 2063

RESULT 15
Q9VEF9 PRELIMINARY; PRT; 1101 AA.
ID Q9VEF9
AC Q9VEF9;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT OCT-2002 (TrEMBLrel. 22, Last annotation update)
DB CG7397 protein.
GN CG7397.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STEAIN=Berkeley.
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard C., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svarksas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
DR EMBL; AB003719; AAF55465.1; -.
DR FlyBase; FBgn0038543; CG7397.
DR InterPro; IPR000219; RhoGEF.
DR Pfam; PF00621; RhoGEF; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS50010; DH 2; 1.
SQ SEQUENCE 1101 AA; 120780 MW; CBIE2D0390EB2F26 CRC64;

Query Match 12.6%; Score 436.5; DB 5; Length 1101;
Best Local Similarity 22.4%; Pred. No. 7e-23;
Matches 140; Conservative 103; Mismatches 196; Indels 187; Gaps 19;

Qy 88 VWILSTLTTSKVLIIDANQGTGVVDTVCNAHVLCISSIPAAASDSYPPGEMFLDSVD 147
Db 596 VVWNCSDGYGVQVCIMSLHPEFNTSCNGVNRILCVASPAYSSA----- 642
Qy 148 NPEDPGADGVLGITVGCATRCNVRPNSCRSGTTPVLDDKQGEVATIANGVNPSQST 207
Db 643 -----ASRNSSEQAGSGQEE-----KDKRHTPQSYT 672
Qy 208 EATEATEVPPDP---GPSEPTATLRPGPLTEHVFTDPAPTSSGPQP-----GSENG 257

Db	673	QQLRDYKRSISPFQSSASATATPEKKLT-----PTTTPVAGGFPADAAATGGSDQT	726
Qy	258	PEPDSSSTRPEPE-----PSGDP-----GAGSSAAPTWWL	288
Db	727	LELNLSSDDTEAAAAASLNVAGSTGLTVERVSPAPSVCHQDSNPEGESNQSTNWI	786
Qy	289	GAQNGWLYVHSAVANWKCLHSIKLK--DSVLSLVHVKGRVLVALADGTTLAIHFRGEDGQ	346
Db	787	GTEGCHTVYNSTNDIRIKKNRIKIEHSAVYSILYLDNRVFSLANGDICIYVLRDGA TS	846
Qy	347	WDLSNYHMLDGLPHHSIRCMVYVDVWVCYKKNVHVIQKTM-----QTEKSPDAPRR	402
Db	847	WNTCSSHCLSIGTWTSPVKNLLNVNGLWCISQIGIIVLDVETLTIVNQIQISSDKP--	904
Qy	403	ESQVRQLAWIGDGWVYSIRLDSTLRLXHAHTHOHLQDVIDTEPYVYSKMLGT-----GKLG F	457
Db	905	---ITNMTVSNHVWISIQNSAHIKCFHSNTHQLVTEVNLAPAVKNMLSNCDEIIRHKA	961
Qy	458	SFRVITALLVAGSLWTGTONGVVISIPLTETVVLHGRQLGLFRANKTSPTSGEGARPGG	517
Db	962	ACLRVTSLLCCOLLITGTSAGVLLTIP-----AQGYEKGA	997
Qy	518	IIHVYGDSDSRAASSPIYCSMAQOLCEGHREDAYKF--FYSVPGNVLATLNGSVILDS	575
Db	998	I-----NIVPTGIP-----HGHTGHRFTFTVETTG-----LEG	1026
Qy	576	PAECPGPAAPASEVEG-----QKLRNVLSGGEYIDFRIGDGEDDETE	620
Db	1027	AASGAGGRDAGSNSDEYTKQGSIKHSKSKSETNNTLIISGGDGIEDFR-----N	1077
Qy	621	EGAGDMSQVPLVLSKAERSHIIWQV	646
Db	1078	SGANSLSEIAG--REDSTNHLHWQI	1101

Search completed: August 23, 2004, 11:14:22
Job time : 135 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 00:01:48 ; Search time 1272 Seconds
(without alignments)
12337.141 Million cell updates/sec

Title: US-10-019-495-26

Perfect score: 3694

Sequence: 1 ctggcaggtctctagtgcg.....gtcgtataaaaaaaaaa 3694

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2003as.*

8: geneseqn2003bs.*

9: geneseqn2003cs.*

10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3694	100.0	3694	3 AAC64282	Aac64282 Human mem
2	3496.4	94.7	4837	4 AAI59280	Aai59280 Human pol
3	3494.2	94.6	4737	3 AAC76926	Aac76926 Human ORF
4	3484.6	94.3	4701	6 ABK09789	Abk09789 Human ova
5	3374	91.3	4956	4 AAI61066	Aai61066 Human pol
6	3182.2	86.1	5007	5 AAs74091	Aas74091 DNA encod
7	1531.6	41.5	4173	3 AAA62899	Aaa62899 Murine JN
8	1531.6	41.5	4200	3 AAA62900	Aaa62900 Murine JN
9	1531.6	41.5	4266	3 AAA62902	Aaa62902 Murine JN
10	1531.6	41.5	4269	3 AAA62901	Aaa62901 Murine JN
11	1457.4	39.5	2431	5 AAs74089	Aas74089 DNA encod
12	486.2	13.2	4688	4 ABA09028	Abao9028 Human sec
13	477	12.9	4286	7 ABX63549	Abx63549 Human cdn
14	442.2	12.0	4667	8 ADA45191	Ada45191 Murine JL
15	413.4	11.2	4675	8 ADA45189	Ada45189 Human JLP
16	382.4	10.4	403	2 AAQ39848	Aaq39848 Expressed
17	382.4	10.4	403	2 AAQ59260	Aaq59260 Human bra
18	370.2	10.0	3113	2 AAV46322	Aav46322 Human sec
19	370.2	10.0	3113	5 AAF98489	Aaf98489 Human cdn
20	360.6	9.8	3791	6 AAs94748	Aas94748 Human imm
21	352.4	9.5	7267	4 AAK69457	Aak69457 Human imm
22	294.4	8.0	3588	4 ABL13755	Ab113755 Drosophil
23	279	7.6	1966	4 AAh16524	Aah16524 Human cdn

24	269.2	7.3	782	5 AAs74086	Aas74086 DNA encod
25	269.2	7.3	1779	5 AAs79794	Aas79794 DNA encod
26	244	6.6	633	3 AAh07478	Aah07478 Human cdn
27	242.4	6.6	1413	3 AAC98939	Aac98939 Human pan
28	237	6.4	237	3 AAC00302	Aac00302 Human sec
29	226	6.1	554	6 ABL90245	Ab190245 Human pol
30	220.4	6.0	10011	4 ABL13754	Ab113754 Drosophil
C 31	219.4	5.9	675	6 ABK63058	Abk63058 Rat seque
C 32	219.4	5.9	675	9 ADB51737	Adb51737 Primary r
33	209.4	5.7	478	5 AAs74070	Aas74070 DNA encod
34	194.8	5.3	633	5 AAs84890	Aas84890 DNA encod
35	194.8	5.3	633	5 AAs74085	Aas74085 DNA encod
36	194.8	5.3	633	5 AAs87991	Aas87991 DNA encod
37	185.2	5.0	2937	4 AAh18341	Aah18341 Human cdn
38	122.4	3.3	128	2 AAT20322	Aat20322 Human gen
39	116.6	3.2	6289	6 ABL65906	Ab165906 Lung canc
40	116.6	3.2	6289	6 ABL65509	Ab165509 Lung canc
41	116.6	3.2	6289	6 ABK64385	Abk64385 Human ben
42	116.6	3.2	7540	6 ABL92090	Ab192090 Human Tum
43	116.6	3.2	7540	7 ABX72015	Abx72015 DNA encod
44	116.6	3.2	7768	7 AET16473	Abt16473 Human int
C 45	108.4	2.9	731	5 ABV30065	Abv30065 Human pro

ALIGNMENTS

RESULT 1

ID	AAC64282	standard; cDNA; 3694 BP.
XX		
AC	AAC64282;	
XX		
DT	23-FEB-2001	(first entry)
XX		
DE	Human membrane-associated protein HUMAP-9 CDNA.	
XX		
KW	Human membrane-associated protein; HUMAP; transgenic organism;	
KW	drug screening; cell signalling modulator; agonist; antagonist;	
KW	cell differentiation modulator; cell proliferation modulator;	
KW	cell proliferative disorder; cancer; cell differentiation disorder;	
KW	developmental disorder; cell signalling disorders; endocrine disorder;	
KW	hyperpituitarism; hypothyroidism; hyperparathyroidism; infection;	
KW	pancreatic disorder; diabetes mellitus; immunological disorder;	
KW	hereditary neuropathy; gonadal steroid hormone associated disorder;	
KW	infertility; ss.	
OS	Homo sapiens.	
XX		
PN	WO200065054-A2.	
XX		
PD	02-NOV-2000.	
XX		
PF	20-APR-2000; 2000WO-US010884.	
XX		
PR	23-APR-1999; 99US-0130694P.	
PR	23-JUN-1999; 99US-0140580P.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
PI	Hillman JI, Bandman O, Tang YT, Lal P, Yue H, Reddy R;	
PI	Arimzai Y, Baughn MR;	
XX		
DR	WPI; 2000-687346/67.	
DR	P-PSDB; AAB29652.	
XX		
PT	Human membrane-associated protein, useful for diagnosis and treatment of	
PT	cell signaling, cell differentiation and cell proliferation disorders	
PT	such as cancer, and for identifying agonists and antagonists.	
XX		
PS	Claim 4; Page 93; 99pp; English.	
XX		
CC	The invention relates to 17 human membrane-associated proteins, HUMAP-1	

to HUMAP-17 (AAB29644-B29660) and the cDNAs encoding them (AAC64274-C64290). The invention also relates to expression constructs, host cells and transgenic organisms comprising a HUMAP nucleic acid sequence; the recombinant preparation of a HUMAP; methods of screening compounds for their ability to modulate HUMAP activity or expression; and pharmaceutical compositions comprising a HUMAP protein, a HUMAP agonist or HUMAP antagonist. The HUMAPs acts as modulators of cell signalling, differentiation and proliferation. A HUMAP is useful for screening a compound for effectiveness as an agonist or antagonist of HUMAP activity. The protein, or the identified agonist or antagonist is useful for treating a disease or condition associated with decreased or increased expression of functional HUMAP. A HUMAP nucleic acid is useful for screening a compound for its ability to alter expression of that particular HUMAP gene. A wide variety of disease may be treated using compositions of the invention. These diseases include cell proliferative disorders (e.g., actinic keratosis, arteriosclerosis); cancer (e.g., breast, bladder, bone marrow, brain and uterine cancer); cell differentiation disorders, in particular developmental disorders (e.g., renal tubular acidosis, anaemia, Cushing's syndrome, achondroplasia, epilepsy, and muscular dystrophy); cell signalling disorders, in particular endocrine disorders such as hypothalamus and pituitary disorders resulting from lesions such as thrombosis; disorders associated with hyperpituitarism (e.g., acromegaly); disorders associated with hypothyroidism (e.g., goitre); hyperparathyroidism; pancreatic disorders such as type I or type II diabetes mellitus; infections; immunological disorders; hereditary neuropathies (e.g., neurofibromatosis); and disorders associated with gonadal steroid hormones (e.g., infertility, endometriosis, polycystic ovary syndrome, osteoporosis, Leydig cell deficiency and gynecomastia). Antibodies which specifically bind HUMAP may be used for the diagnosis of disorders associated with the expression of HUMAP, or in assays to monitor patients being treated with HUMAP or agonists, antagonists or inhibitors of HUMAP. The present sequence represents a HUMAP cDNA of the invention

Sequence 3694 BP: 677 A; 1265 C; 1169 G; 583 T; 0 U; 0 Other;

[illegible]

XX AC AAI59280;
XX 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 1483.
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokineic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia; ss.
XX Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US034263.
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-00488725.
XX 25-APR-2000; 2000US-00552317.
XX 20-JUN-2000; 2000US-00598042.
XX 19-JUL-2000; 2000US-00620312.
XX 03-AUG-2000; 2000US-00653450.
XX 14-SEP-2000; 2000US-00662191.
XX 19-OCT-2000; 2000US-00693036.
XX 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX P-PSDB; AAM40124.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX Claim 1; SEQ ID NO 1483; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
XX encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
SQ Sequence 4837 BP; 978 A; 1600 C; 1507 G; 752 T; 0 U; 0 Other;
Query Match 94.7%; Score 3496.4; DB 4; Length 4837;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3503; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 180 GCGTCCTGCTCCCTGACCTGAGTCCCAACGGGGGCCAGGAGGACACCGGATGAAGAA 239
DB 1324 GCCCGCCAAAGTACAAGCAGCTGAGTCCCAACGGGGGCCAGGAGGACACCGGATGAAGAA 1383
QY 240 CGTGCCGGTGCCGGTGTACTGCGGCCCTCTGTGGAGAGGAGGCCCAACCATGAAGCTGTG 299

Db 1384 CGTGCCGGTGCCGGTGTACTGCGGCCCTCTGGTGGAGAGGACCCACCATGAAGCTGTG 1443
QY 300 GTGTGCCCGGGGCGTCAACTGAGCGGGTGGAGGCCCAATGAGAGCAACGCTGGGAATGG 359
Db 1444 GTGTGCCCGGGGCGTCAACTGAGCGGGTGGAGGCCCAATGAGAGCAACGCTGGGAATGG 1503
QY 360 AGTCAAGCCAGCCAGCCAGCCGCGATCCCTGACCTGCGACCGCGAGCGGAGAGCGGAGCC 419
Db 1504 AGTCAAGCCAGCCAGCCAGCCGCGATCCCTGACCTGCGACCGCGAGCGGAGAGCGGAGCC 1563
QY 420 CAACAGCGCCACACGCTCTCCCGAGAGAGAAAGGCGCAAGGAGCTCCCTGAAATGGAGCG 479
Db 1564 CAACAGCGCCACACGCTCTCCCGAGAGAGAGAGCGCAAGGAGCTCCCTGAAATGGAGCG 1623
QY 480 CACCTCCAGCCGGGTGTGGATCTTGACAGCACCCCTGACCAACAGAGGTGGTGTATCAT 539
Db 1624 CACCTCCAGCCGGGTGTGGATCTTGACAGCACCCCTGACCAACAGAGGTGGTGTATCAT 1683
QY 540 CGAGCCCAACAGCCGGGCGAGGGTGGAGCCAGTTCACCGTCTGCAACGCGCAGCTGCT 599
Db 1684 CGAGCCCAACAGCCGGGCGAGGGTGGAGCCAGTTCACCGTCTGCAACGCGCAGCTGCT 1743
QY 600 GTGCATCTCCAGCATCCCGCGGCGAGCGACAGCGACTTACCCCTCCCGGGAGATGTTCT 659
Db 1744 GTGCATCTCCAGCATCCCGCGGCGAGCGACAGCGACTTACCCCTCCCGGGAGATGTTCT 1803
QY 660 GGAACAGCGAAGTGAACCCAGAGAGCCCGGGCGAGAGTGGCGGTGTATCACCT 719
Db 1804 GGAACAGCGAAGTGAACCCAGAGAGCCCGGGCGAGAGTGGCGGTGTATCACCT 1863
QY 720 GGTGGGCTGTGCCACCCGCTGCAACGTCGCCGCGAGCAACTGCTCTCCCGAGGGGACAC 779
Db 1864 GGTGGGCTGTGCCACCCGCTGCAACGTCGCCGCGAGCAACTGCTCTCCCGAGGGGACAC 1923
QY 780 CCCAGTGTAGCAAGGGGCGAGGGGAGTGGCCACCATCCCAACGGGAGGTCAACCC 839
Db 1924 CCCAGTGTAGCAAGGGGCGAGGGGAGTGGCCACCATCCCAACGGGAGGTCAACCC 1983
QY 840 GTCCAGTCCACAGAGAGGGCCACAGAGCCACCGAGGTGCCAGACCTGGGCGCCAGCA 899
Db 1984 GTCCAGTCCACAGAGAGGGCCACAGAGCCACCGAGGTGCCAGACCTGGGCGCCAGCA 2043
QY 900 GCCAGAGACAGCCACATTGGCGGCCCGGCGCTCTACAGAGCACGCTTCTCACTGACCCAGC 959
Db 2044 GCCAGAGACAGCCACATTGGCGGCCCGGCGCTCTACAGAGCACGCTTCTCACTGACCCAGC 2103
QY 960 CCCGACCCCGCTCTCTGGCCCCCAGCCTGGCAGCGAGAACCGGGCCAGAGCCTTGACAGCAG 1019
Db 2104 CCCGACCCCGCTCTCTGGCCCCCAGCCTGGCAGCGAGAACCGGGCCAGAGCCTTGACAGCAG 2163
QY 1020 GAGCACCGCCAGAGCCAGAGCCCGGGGAGCCCGAGGGAGCGAGGAGCGAGGTGCTGC 1079
Db 2164 GAGCACCGCCAGAGCCAGAGCCCGGGGAGCCCGAGGGAGCGAGGAGCGAGGTGCTGC 2223
QY 1080 ACCCACCATTGGCTGGGAGCCCAAGACGGCTGCTCTATGTGCACTCGGCTGTGGCCAA 1139
Db 2224 ACCCACCATTGGCTGGGAGCCCAAGACGGCTGCTCTATGTGCACTCGGCTGTGGCCAA 2283
QY 1140 CTGGAAGAGTGCCTGCACTCCATCAAGCTGAAGGATCTGTGTGAGCGCTGGTGATGT 1199
Db 2284 CTGGAAGAGTGCCTGCACTCCATCAAGCTGAAGGATCTGTGTGAGCGCTGGTGATGT 2343
QY 1200 CAAAGGCGGTGTGCTGGTGGCTCTGCGGAGCGGAGCCCTGGCCATCTTCCACCGGTGA 1259
Db 2344 CAAAGGCGGTGTGCTGGTGGCTCTGCGGAGCGGAGCCCTGGCCATCTTCCACCGGTGA 2403
QY 1260 AGATGGCCAGTGGGATCTGAGCAACTATCACCTAATGGACCTGGGCGCACCCGACACATC 1319
Db 2404 AGATGGCCAGTGGGATCTGAGCAACTATCACCTAATGGACCTGGGCGCACCCGACACATC 2463
QY 1320 CATCGCTGCAATGGCTGTGTGTGTAAGACCGCTGTGTGTGCTACAAAGACAGGTGCA 1379

Db	2464	CATCCGCTGCATGGCTGTGTGTGTA	CGACCGCGTGTGTGTGCTACAAGAA	CAAGGTGCA	2523
QY	1380	CGTCAATCCAGCCCAAGAACCATG	CAGATAGAGAAAGTCAATTTTGA	CGCCCAACCCGCGCGGA	1439
Db	2524	CGTCAATCCAGCCCAAGAACCATG	CAGATAGAGAAAGTCAATTTTGA	CGCCCAACCCGCGCGGA	2583
QY	1440	GAGCCAGGTGGCGGAGCTGGCGT	GTGGATGGCGGATGGCGATGGCG	ATGGGTATGGGTATGGGTATGGGT	1499
Db	2584	GAGCCAGGTGGCGGAGCTGGCGT	GTGGATGGCGGATGGCGATGGCG	ATGGGTATGGGTATGGGTATGGGT	2643
QY	1500	CTCCACCTTGAGGCTCTACCATG	CAACAGGACAGCATCTACAGGA	CGTGGACATGGA	1559
Db	2644	CTCCACCTTGAGGCTCTACCATG	CAACAGGACAGCATCTACAGGA	CGTGGACATGGA	2703
QY	1560	GCCCTACGTCAGCAAGATCGTAG	GCACCTGGCACTGGCAAGCTGGG	TTCTCCTTCGTACGATCAC	1619
Db	2704	GCCCTACGTCAGCAAGATCGTAG	GCACCTGGCACTGGCAAGCTGGG	TTCTCCTTCGTACGATCAC	2763
QY	1620	GGCCCTGCTGTGGCGGAGCCGGT	CTGGGTGGGCACCGGCAACGG	AGTGGTCAATCTC	1679
Db	2764	GGCCCTGCTGTGGCGGAGCCGGT	CTGGGTGGGCACCGGCAACGG	AGTGGTCAATCTC	2823
QY	1680	CATCCCCCTGACAGAGACTGGT	CTGCACCGAGGCGAGCTCTGG	GGGCTCCGAGCCAA	1739
Db	2824	CATCCCCCTGACAGAGACTGGT	CTGCACCGAGGCGAGCTCTGG	GGGCTCCGAGCCAA	2883
QY	1740	TAAGACATCCCCCACTCTGGG	AGGGCGCCCGTCTGGGTGGGC	ACCGGCAACGGAGTGGTCAATCTC	1799
Db	2884	TAAGACATCCCCCACTCTGGG	AGGGCGCCCGTCTGGGTGGGC	ACCGGCAACGGAGTGGTCAATCTC	2943
QY	1800	CGATGACAGAGTGACAGGGG	CGCAGCAGTTCATCCCTACTG	CTCCATGGCCCAAGGC	1859
Db	2944	CGATGACAGAGTGACAGGGG	CGCAGCAGTTCATCCCTACTG	CTCCATGGCCCAAGGC	3003
QY	1860	CCAGCTATGCTTCCATGGG	CAACCGGATGCGGTGAAGTTC	TTTGTCTGGTGCCAGGAA	1919
Db	3004	CCAGCTATGCTTCCATGGG	CAACCGGATGCGGTGAAGTTC	TTTGTCTGGTGCCAGGAA	3063
QY	1920	CGTGTGGCCCACTCTGAATGG	CAGTGTGCTGGACAGCCCGAG	CGCGGCGCCCTGGGCCAGC	1979
Db	3064	CGTGTGGCCCACTCTGAATGG	CAGTGTGCTGGACAGCCCGAG	CGCGGCGCCCTGGGCCAGC	3123
QY	1980	TGCCCCCTGCTCGGAGTGC	AGGGCCAGAACTGCTGGTGC	TGAGCGGCGG	2039
Db	3124	TGCCCCCTGCTCGGAGTGC	AGGGCCAGAACTGCTGGTGC	TGAGCGGCGG	3183
QY	2040	GGAGGGCTACATCGACTTCC	CAATGGAGACCGAGACAGCA	CGAGAGAGGGCGC	2099
Db	3184	GGAGGGCTACATCGACTTCC	CAATGGAGACCGAGACAGCA	CGAGAGAGGGCGC	3243
QY	2100	AGGGGACATGAGCCAGGTGA	AGCCCGTGTCTGTCCAAGG	ACAGCGCAGTCAATCATCGT	2159
Db	3244	AGGGGACATGAGCCAGGTGA	AGCCCGTGTGTCTCAAGG	ACAGCGCAGTCAATCATCGT	3303
QY	2160	GTGGCAGGTGTCATACCC	CCGAGTGAAGTGTCTGCC	TGCCCCGACCTGTACAT	2219
Db	3304	GTGGCAGGTGTCATACCC	CCGAGTGAAGTGTCTGCC	TGCCCCGACCTGTACAT	3363
QY	2220	AGGACCCCGACACCTGAC	CCCCCGCCCGCGCGGGGT	AGCAGCAGCGCGCGCGC	2279
Db	3364	AGGACCCCGACACCTGAC	CCCCCGCCCGCGCGGGGT	AGCAGCAGCGCGCGCGC	3423
QY	2280	CCCTCTTTCACTCTCAACT	GTGAGCTTTTCACTGTGG	CCCTTCAGCGGGGAGG	2339
Db	3424	CCCTCTTTCACTCTCAACT	GTGAGCTTTTCACTGTGG	CCCTTCAGCGGGGAGG	3483
QY	2340	GAGTGGGGGATGCGGAT	CAAGTGGGAGGAGGGAGGG	GGTGTCTTCCACCCGAGGGA	2399
Db	3484	GAGTGGGGGATGCGGAT	CAAGTGGGAGGAGGGAGGG	GGTGTCTTCCACCCGAGGGA	3543
QY	2400	AGATGCTCTCGGACAGT	TTCCCGGGGAGCTCTCGG	CGAGCTTCCAGCCAGAGTCTCA	2459
Db	3544	AGATGCTCTCGGACAGT	TTCCCGGGGAGCTCTCGG	CGAGCTTCCAGCCAGAGTCTCA	3603
QY	2460	AGTCCAGGGACCTTGGG	CCAGCGCAGGAGAAATCCG	AGGTGGTCTTGCTTACCCCTG	2519
Db	3604	AGTCCAGGGACCTTGGG	CCAGCGCAGGAGAAATCCG	AGGTGGTCTTGCTTACCCCTG	3663
QY	2520	GGCTCTACTCCCAAGCAC	CCCTTGAGGAGCAGGGGCT	CCCCCGCGCGGAGGCTGCCT	2579
Db	3664	GGCTCTACTCCCAAGCAC	CCCTTGAGGAGCAGGGGCT	CCCCCGCGCGGAGGCTGCCT	3723
QY	2580	GGCTTGGGCCCCACTCTG	CATGTGCTCATGGGGCCAC	CGCTCTCTGGGGCCCTCACTC	2639
Db	3724	GGCTTGGGCCCCACTCTG	CATGTGCTCATGGGGCCAC	CGCTCTCTGGGGCCCTCACTC	3783
QY	2640	TGCTTAGGGAGCTGGG	CCAGGCACTAGCCTTTGGCC	AGGGAGTGGGCTCAGGCTGCC	2699
Db	3784	TGCTTAGGGAGCTGGG	CCAGGCACTAGCCTTTGGCC	AGGGAGTGGGCTCAGGCTGCC	3843
QY	2700	CAGGTGCTGACCCAGC	CCGCTCTCTGAGGGCTCC	CGCTCTCAAGCCCTTATCTT	2759
Db	3844	CAGGTGCTGACCCAGC	CCGCTCTCTGAGGGCTCC	CGCTCTCAAGCCCTTATCTT	3903
QY	2760	GTCTGTCCCCACCC	CAGCTGTCCCTGGCCAGG	AGCTGGGCATATAAGCACGAG	2819
Db	3904	GTCTGTCCCCACCC	CAGCTGTCCCTGGCCAGG	AGCTGGGCATATAAGCACGAG	3963
QY	2820	CTCCCTGGGCGAGCTGT	TGAGAACAGACTGTAC	CCCCATCTCTGCCCATCAGG	2879
Db	3964	CTCCCTGGGCGAGCTGT	TGAGAACAGACTGTAC	CCCCATCTCTGCCCATCAGG	4023
QY	2880	CTCTTGGCAGCCCGT	CTGACCCGTGCCCCCAGG	CTCTCTGGGCGAGAGACTCAC	2939
Db	4024	CTCTTGGCAGCCCGT	CTGACCCGTGCCCCCAGG	CTCTCTGGGCGAGAGACTCAC	4083
QY	2940	CTTGAGGAGTGGG	CCCTGTGTCTCCAGAA	AGCCGCCCTGCTGGGCGAGAG	3059
Db	4084	CTTGAGGAGTGGG	CCCTGTGTCTCCAGAA	AGCCGCCCTGCTGGGCGAGAG	4143
QY	3000	AGGCTGCCAGGGC	AGGCGCTCAGAGAGG	AGGCGGCGCTCTCCGGGAT	4203
Db	4144	AGGCTGCCAGGGC	AGGCGCTCAGAGAGG	AGGCGGCGCTCTCCGGGAT	4263
QY	3060	CAGTCTTAGGACACAGG	CTCAGCTCAGCTTACTA	AGTGCCTTAGGCTT	3179
Db	4204	CAGTCTTAGGACACAGG	CTCAGCTCAGCTTACTA	AGTGCCTTAGGCTT	4323
QY	3120	CCTCTTGCACGGGCT	CCACGGAGCCAGCTCC	CAGACACGCTTACTAGG	3179
Db	4264	CCTCTTGCACGGGCT	CCACGGAGCCAGCTCC	CAGACACGCTTACTAGG	4323
QY	3180	GCCCGCTGTGGCT	GTCTCCAGGAGCAACAG	AGGAGCCACCAAGGAGG	3239
Db	4324	GCCCGCTGTGGCT	GTCTCCAGGAGCAACAG	AGGAGCCACCAAGGAGG	4383
QY	3240	CTGAGGATGAGCG	CGCCCGAGCGGCTCC	CAAGCCCGAGAGGCGAC	3299
Db	4384	CTGAGGATGAGCG	CGCCCGAGCGGCTCC	CAAGCCCGAGAGGCGAC	4443
QY	3300	CTGCTCTCCCTG	CCAGCTGTGGGCTCT	CTGCTTCTTCTTCTTCT	3359
Db	4444	CTGCTCTCCCTG	CCAGCTGTGGGCTCT	CTGCTTCTTCTTCTTCT	4503
QY	3360	TCTGTCTGGAGG	CGCCCTTATGAGGG	CAGCCCGAGCCACCA	3419
Db	4504	TCTGTCTGGAGG	CGCCCTTATGAGGG	CAGCCCGAGCCACCA	4563
QY	3420	AGCACAGATCTTGGG	AGCTGCCCAAGCCCGCT	TGGCCACCGAGGGCTG	3479
Db	4564	AGCACAGATCTTGGG	AGCTGCCCAAGCCCGCT	TGGCCACCGAGGGCTG	4623
QY	3480	GGCTGCGGGCTTCT	CCCCACCACTTCC	CGAGCTTGTATGTC	3539
Db	4624	GGCTGCGGGCTTCT	CCCCACCACTTCC	CGAGCTTGTATGTC	4683

QY 3540 GTCTGTTCCCCAGGATCTCGAAGTGAATCGCGGCTGAGCAGTGGGGGGCTGGGGAGG 3599
D 4684 GTCTGTTCCCCAGGATCTCGAAGTGAATCGCGGCTGAGCAGTGGGGGGCTGGGGAGG 4743
QY 3600 GGTGACGATTCCTCTAGGCTTTGGCCCTGCAAGCAAAACCCACATATCTGCTCTGTATGT 3659
D 4744 GGTGACGATTCCTCTAGGCTTTGGCCCTGCAAGCAAAACCCACATATCTGCTCTGTATGT 4803
QY 3660 AATAAATGCTTAACTGCTGTAATAAATAAATAA 3693
D 4804 AATAAATGCTTAACTGCTGTAATAAATAAATAA 4837

RESULT 3
AAC76926
ID AAC76926 standard; cDNA; 4737 BP.

AC AAC76926;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2481 polynucleotide sequence SEQ ID NO:4961.

Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
anticongulant; osteopathic; antiarthritic; immunosuppressant; cardiant;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; antiinflammatory;
antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
antianemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antiinflammatory disease; coagulation;
thrombosis; contraceptive; ss.

OS Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX 31-MAR-1999; 99US-0127607P.

XX 02-APR-1999; 99US-0127636P.

XX 05-APR-1999; 99US-0127728P.

XX 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB42717.

XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX Claim 5; Page 4128-4131; 5507pp; English.
XX AAC7446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticongulant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
CC sequences can be used for determining the presence of or predisposition

CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 4737 BP; 966 A; 1574 C; 1470 G; 727 T; 0 U; 0 Other;

Query Match 94.6%; Score 3494.2; DB 3; Length 4737;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 3502; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 180 GCCTCCCTGCTCCCTGAGCTGAGTCCCAACGGGGGCCAGGAGACACCGGATGAAGAA 239
D 1203 GCCCGCCCAAGTACAAAGCAGCTGAGTCCCAACGGGGGCCAGGAGACACCGGATGAAGAA 1262
QY 240 CGTCCGGTGGCGGTGACTGCGCCCTCTCGTGGAGAAGACCCACCATGAAGCTGTG 299
D 1263 CGTCCGGTGGCGGTGACTGCGCCCTCTCGTGGAGAAGACCCACCATGAAGCTGTG 1322
QY 300 GTGTCCCGCGGGCTCAACCTGAGCGGTGGAGCCCAATGAGGACGACGCTGGGAATGG 359
D 1323 GTGTCCCGCGGGCTCAACCTGAGCGGTGGAGCCCAATGAGGACGACGCTGGGAATGG 1382
QY 360 AGTCAAGCCAGCGCCGCGCGATCCCTCACTGCGACCGCGAAGGAGACGGGAGCC 419
D 1383 AGTCAAGCCAGCGCCGCGCGATCCCTCACTGCGACCGCGAAGGAGACGGGAGCC 1442
QY 420 CAAGAGCGCCACACGCTTCCCGAGAAGAACGAGCCAGGAGCTCCCTGAATGGAGCG 479
D 1443 CAAGAGCGCCACACGCTTCCCGAGAAGAACGAGCCAGGAGCTCCCTGAATGGAGCG 1502
QY 480 CACCTCCAGCGGGTGTGATCCTGACGAGACACCTGACCAACAGGAGTGGTATCAT 539
D 1503 CACCTCCAGCGGGTGTGATCCTGACGAGACACCTGACCAACAGGAGTGGTATCAT 1562
QY 540 CGACGCCAACAGCGCGGCGACGGTGGTGGACCGAGTTCACCGTCTGCAACGCGACGTGT 599
D 1563 CGACGCCAACAGCGCGGCGACGGTGGTGGACCGAGTTCACCGTCTGCAACGCGACGTGT 1622
QY 600 GTGCATCTCCAGCATCCCGCGGCGAGCAGCAGCAGCTACCTCCCGGGGAGTGTCTCT 659
D 1623 GTGCATCTCCAGCATCCCGCGGCGAGCAGCAGCAGCTACCTCCCGGGGAGTGTCTCT 1682
QY 660 GGACAGCGACGTAACCCAGAGGACCCCGGCGCAGATGGCGTGTCTGGCGGTATCACCTT 719
D 1683 GGACAGCGACGTAACCCAGAGGACCCCGGCGCAGATGGCGTGTCTGGCGGTATCACCTT 1742
QY 720 GGTGGGCTGTGCCACCCCGCTGCAACGTCGCGCGAGCAACTGCTCTCTCCGAGGGGACAC 779
D 1743 GGTGGGCTGTGCCACCCCGCTGCAACGTCGCGCGAGCAACTGCTCTCTCCGAGGGGACAC 1802
QY 780 CCCAGTGTAGCAAGGGGCGAGGGAGTGGCCACCATCTCCCAAGGGAGGTCAACCC 839
D 1803 CCCAGTGTAGCAAGGGGCGAGGGAGTGGCCACCATCTCCCAAGGGAGGTCAACCC 1862
QY 840 GTCCCAAGTCCACAGAGGAGGCCACAGAGGCGACGAGGTGCGAGACCTCTGGGCCACGGA 899
D 1863 GTCCCAAGTCCACAGAGGAGGCCACAGAGGCGACGAGGTGCGAGACCTCTGGGCCACGGA 1922
QY 900 GCCAGAGACAGCCACATTCGCGCCCGGGCTCTTCACAGAGCAGTCTTCACTGACCCAGC 959
D 1923 GCCAGAGACAGCCACATTCGCGCCCGGGCTCTTCACAGAGCAGTCTTCACTGACCCAGC 1982
QY 960 CCGGACCCCGTCTCTGGCCCCCGCCGCTGCGAGGAGACGGGCCAGAGCCTCAGAGCAG 1019
D 1983 CCGGACCCCGTCTCTGGCCCCCGCCGCTGCGAGGAGACGGGCCAGAGCCTCAGAGCAG 2042

QY	1020	CAGCACACGGCCAGAGCCAGAGCCAGCGGGACCCCAACGGAGAGGAGCAGAGTGTGC	1079
Db	2043	CAGCACACGGCCAGAGCCAGAGCCAGCGGGACCCCAACGGAGAGGAGCAGAGTGTGC	2102
QY	1080	ACCCACCAATGTGGCTGGAGGCCAGAAACGGCTGGCTCTAATGTGCACTGGCTGTGGCCAA	1139
Db	2103	ACCCACCAATGTGGCTGGAGGCCAGAAACGGCTGGCTCTAATGTGCACTGGCTGTGGCCAA	2162
QY	1140	CTGGAAGAAGTGTGGCTGTGGCTCTGGCGGACGGGACCCCTGGCCATCTTCCACCGTGGTGA	1199
Db	2163	CTGGAAGAAGTGTGGCTGTGGCTCTGGCGGACGGGACCCCTGGCCATCTTCCACCGTGGTGA	2222
QY	1200	CAAAAGCCGTGTGGCTGTGGCTCTGGCGGACGGGACCCCTGGCCATCTTCCACCGTGGTGA	1259
Db	2223	CAAAAGCCGTGTGGCTGTGGCTCTGGCGGACGGGACCCCTGGCCATCTTCCACCGTGGTGA	2282
QY	1260	AGATGGCCAGTGGGATCTGAGCAATATCACTAATAGCACTTGGGCCACCCGCCACCACTC	1319
Db	2283	AGATGGCCAGTGGGATCTGAGCAATATCACTAATAGCACTTGGGCCACCCGCCACCACTC	2342
QY	1320	CATCCGCTGCATGGCTGTGGTGTAGCACCGCGTGTGGTGTGGCTACAGAAACAGGTGCA	1379
Db	2343	CATCCGCTGCATGGCTGTGGTGTAGCACCGCGTGTGGTGTGGCTACAGAAACAGGTGCA	2402
QY	1380	CGTCATCCAGCCCAAGACCATGCAATAGAGAGTCAATTTGAGCGCCACCCCGCGCGGGA	1439
Db	2403	CGTCATCCAGCCCAAGACCATGCAATAGAGAGTCAATTTGAGCGCCACCCCGCGCGGGA	2462
QY	1440	GAGCCAGTGGCGAGCTGGCGTGGATCGGCGATGGCGTATGGGTGTCCATCCGCTTGA	1499
Db	2463	GAGCCAGTGGCGAGCTGGCGTGGATCGGCGATGGCGTATGGGTGTCCATCCGCTTGA	2522
QY	1500	CTCCACCTGAGGCTCTACCATGACACACACACACAGCATCTACAGACGTGGACATGGA	1559
Db	2523	CTCCACCTGAGGCTCTACCATGACACACACACACAGCATCTACAGACGTGGACATGGA	2582
QY	1560	GCCCTACGTACAGAGATGCTAGGCACCTGGCAAGCTGGGTCTTCTCTCGTACGCATCAC	1619
Db	2583	GCCCTACGTACAGAGATGCTAGGCACCTGGCAAGCTGGGTCTTCTCTCGTACGCATCAC	2642
QY	1620	GGCCCTGTGTGCGGGGACCGCGCTCTGGGTGGGACCGGCAACGGAGTGTCACTC	1679
Db	2643	GGCCCTGTGTGCGGGGACCGCGCTCTGGGTGGGACCGGCAACGGAGTGTCACTC	2702
QY	1680	CATCCCTCTGACAGAGTGTGTCTGACCGAGGCGCAGCTCTGGGGCTCCAGGCAAA	1739
Db	2703	CATCCCTCTGACAGAGTGTGTCTGACCGAGGCGCAGCTCTGGGGCTCCAGGCAAA	2762
QY	1740	TAGACATCCCCACCTCTGGGAGGGCGCGCGTCCCGGGGGCATCATCCAGTGTATGG	1799
Db	2763	TAGACATCCCCACCTCTGGGAGGGCGCGCGTCCCGGGGGCATCATCCAGTGTATGG	2822
QY	1800	CGATGACAGAGTGAACAGGGGGCCAGAGCTTCACTCCCTTACTGTCTCCATGGCCCAAGC	1859
Db	2823	CGATGACAGAGTGAACAGGGGGCCAGAGCTTCACTCCCTTACTGTCTCCATGGCCCAAGC	2882
QY	1860	CCAGTATGCTTCCATGCGCACCGCGATGCGTGAAGTCTTTGTCTCGTGGTCCAGGGAA	1919
Db	2883	CCAGTATGCTTCCATGCGCACCGCGATGCGTGAAGTCTTTGTCTCGTGGTCCAGGGAA	2942
QY	1920	CGTGTGGCCACCTGAAATGCAAGTGTGGACAGCCCGAGGGCCCTTGGGCCAGC	1979
Db	2943	CGTGTGGCCACCTGAAATGCAAGTGTGGACAGCCCGAGGGCCCTTGGGCCAGC	3002
QY	1980	TGCCCTGTCTCGAGTTCAGAGGCCAGAAAGCTGCGGAAAGTGTGTGTGTGAGCGGCGG	2039
Db	3003	TGCCCTGTCTCGAGTTCAGAGGCCAGAAAGCTGCGGAAAGTGTGTGTGTGAGCGGCGG	3062
QY	2040	GGAGGGCTACATCGACTTCGCAATTGAGAGCGGAGAGGACGACAGACGGAGGGCGC	2099
Db	3063	GGAGGGCTACATCGACTTCGCAATTGAGAGCGGAGAGGACGACAGACGGAGGGCGC	3122

QY	2100	AGGGACATGACGACAGGTGAAGCCCGTGTGTCCAAAGGACAGAGGAGTCAATCATCT	2159
Db	3123	AGGGACATGACGACAGGTGAAGCCCGTGTGTCCAAAGGACAGAGGAGTCAATCATCT	3182
QY	2160	GTGGCAGGTGTCTTACACACCCCGAGTGAAGTGTGTCTGCTTGGCTGGCGAGCTGTACAT	2219
Db	3183	GTGGCAGGTGTCTTACACACCCCGAGTGAAGTGTGTCTGCTTGGCTGGCGAGCTGTACAT	3242
QY	2220	AGGACCCCGGACCACTGACCCCGCGCGCGCGGGGTAGCCAGCCAGGCGCGCGCGC	2279
Db	3243	AGGACCCCGGACCACTGACCCCGCGCGCGCGGGGTAGCCAGCCAGGCGCGCGCGC	3302
QY	2280	CCCTCTTCTAACCCTTCAACCTGAGCTTCACTGTAGTGTGGCCCTTCCAGCGGGCAGG	2339
Db	3303	CCCTCTTCTAACCCTTCAACCTGAGCTTCACTGTAGTGTGGCCCTTCCAGCGGGCAGG	3362
QY	2340	GAGTGGGGGATGGGGATCAGCTGGAGGAGGAGGGGAGGGGTGCTTCCACCCAGGGGA	2399
Db	3363	GAGTGGGGGATGGGGATCAGCTGGAGGAGGAGGGGAGGGGTGCTTCCACCCAGGGGA	3422
QY	2400	AGATGTCTCTGGGACAGTTTCCCGGGCAGCTTCTGGCCAGCTTCCAGCCAGAGTCTCA	2459
Db	3423	AGATGTCTCTGGGACAGTTTCCCGGGCAGCTTCTGGCCAGCTTCCAGCCAGAGTCTCA	3482
QY	2460	AGTCCAGGGACCTTGGGCCCGCAGCGCAGGAGAGATCCGAGTGTCTTGGCTCTACCTG	2519
Db	3483	AGTCCAGGGACCTTGGGCCCGCAGCGCAGGAGAGATCCGAGTGTCTTGGCTCTACCTG	2542
QY	2520	GGCTCTCTTCTTCCCGCAGCACCCCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2579
Db	3543	GGCTCTCTTCTTCCCGCAGCACCCCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3602
QY	2580	GGCTGTGGGCCCACTCTGTATGTCTCATGGGGCCACCTGTGCTTCTGGGGCCCTCACTC	2639
Db	3603	GGCTGTGGGCCCACTCTGTATGTCTCATGGGGCCACCTGTGCTTCTGGGGCCCTCACTC	3662
QY	2640	TGCTTGGGGAGCTGGGCCAGGCACTAGCTTGTGCTTGGCCAGGAGGAGTGGGCTTGGG	2699
Db	3663	TGCTTGGGGAGCTGGGCCAGGCACTAGCTTGTGCTTGGCCAGGAGGAGTGGGCTTGGG	3722
QY	2700	CAGTGTGCTGTGACCCCGCAGCGCGCTTCTTGGGGCCCTCCCGCTGTCAAGCTCTATCT	2759
Db	3723	CAGTGTGCTGTGACCCCGCAGCGCGCTTCTTGGGGCCCTCCCGCTGTCAAGCTCTATCT	3782
QY	2760	GTCTGTCCCAACCCCGCAGCTGTCCCTGCCAGGAGCTGGCATAAAGACAGAGGCCCGG	2819
Db	3783	GTCTGTCCCAACCCCGCAGCTGTCCCTGCCAGGAGCTGGCATAAAGACAGAGGCCCGG	3842
QY	2820	CTCCCTGGGGCAGCTGTGACACAGAGACTGTCTACCCCATCTGCTGCGCATGAGGAGG	2879
Db	3843	CTCCCTGGGGCAGCTGTGACACAGAGACTGTCTACCCCATCTGCTGCGCATGAGGAGG	3902
QY	2880	CTCTTGGCAGCCCGCTTGTGACCCCGTGTCCCGCCAGGCTTGTGCTGGCGGAGAGACTCA	2939
Db	3903	CTCTTGGCAGCCCGCTTGTGACCCCGTGTCCCGCCAGGCTTGTGCTGGCGGAGAGACTCA	3962
QY	2940	CTTGGAGGAGTGGGCCCTTGAAGTCTGTCTCCCTCCAGAAAGCCCGCAGGGTGGGATTTCT	2999
Db	3963	CTTGGAGGAGTGGGCCCTTGAAGTCTGTCTCCCTCCAGAAAGCCCGCAGGGTGGGATTTCT	4022
QY	3000	AGGCTGCCAGGGCAGGCCCGAGGCTCAGGAAGAGGGGAGGGCCCTTGGCTTCTCCGGGAT	3059
Db	4023	AGGCTGCCAGGGCAGGCCCGAGGCTCAGGAAGAGGGGAGGGCCCTTGGCTTCTCCGGGAT	4082
QY	3060	CAGTCTCTAGGACACAGGCTCAGCTCAGGTGTGATGGGGATGATGTCTCCCGGGGCTG	3119
Db	4083	CAGTCTCTAGGACACAGGCTCAGCTCAGGTGTGATGGGGATGATGTCTCCCGGGGCTG	4142
QY	3120	CCTCTCTGACGGGGCTTCCAGGAGGCCAGCTCCAGACAGCTTAAAGTGTCTTAGGGTT	3179
Db	4143	CCTCTCTGACGGGGCTTCCAGGAGGCCAGCTCCAGACAGCTTAAAGTGTCTTAGGGTT	4202
QY	3180	GCCCCGTGTGGCTGTCTCCAGGAGGCAACAGAGAGGCCCAACAGAGAGGCGCCGTGGGG	3239

Db 4203 GCGCGTGTGGCTTCTCCAGGAGCAACAGAGAGGCCACCAAGAGAGGCCCGTGGG 4262
Qy 3240 CTGAGATGAGCGCGCCCGCAGCGACTCCAAAGCCCGCAGAGGCGACAAGCCACCTTGA 3299
Db 4263 CTGAGATGAGCGCGCCCGCAGCGACTCCAAAGCCCGCAGAGGCGACAAGCCACCTTGA 4322
Qy 3300 CTGCTCTCCCTGCGCGAGCTGGGCTCTCTGGCCCTATTCTTACCTTCCAGGCGCACTGCAC 3359
Db 4323 CTGCTCTCCCTGCGCGAGCTGGGCTCTCTGGCCCTATTCTTACCTTCCAGGCGCACTGCAC 4382
Qy 3360 TCCTGTCTGGAGGCGCTTATGAGGCGAGCGCCCGCCCGCAGCCACCCCAACAGAGA 3419
Db 4383 TCCTGTCTGGAGGCGCTTATGAGGCGAGCGCCCGCCCGCAGCCACCCCAACAGAGA 4442
Qy 3420 AGCAGATCTTGGGAGCTGCCCAAGCCCGCTGGCCACCGAGGCTGCGAGCGCT 3479
Db 4443 AGCAGATCTTGGGAGCTGCCCAAGCCCGCTGGCCACCGAGGCTGCGAGCGCT 4502
Qy 3480 GCGTGGCGGCTTCTCCACACCCCTGCGCAGCTCCACTGTGATGTATGTCGGCTCCCTC 3539
Db 4503 GCGTGGCGGCTTCTCCACACCCCTGCGCAGCTCCACTGTGATGTATGTCGGCTCCCTC 4562
Qy 3540 GTCTGTTCCTCCAGGATCTCGAAGTGACTCCGGGCTGAGCAGTGGGGCGGTGGGGAGG 3599
Db 4563 GTCTGTTCCTCCAGGATCTCGAAGTGACTCCGGGCTGAGCAGTGGGGCGGTGGGGAGG 4622
Qy 3600 GGTGACGATCTCTCAGGCTTTGGCCCTGCAAGCAAAACCCACATATCTCTGTATGT 3659
Db 4623 GGTGACGATCTCTCAGGCTTTGGCCCTGCAAGCAAAACCCACATATCTCTGTATGT 4682
Qy 3660 AATAAATCTCTTAAGTCTGTAATAAATAAATAA 3694
Db 4683 AATAAATGCTTAAGTCTGTAATAAATAAATAA 4717

RESULT 4
ABK09789
ID ABK09789 standard; cDNA; 4701 BP.
AC ABK09789;
DT 14-MAR-2002 (first entry)
XX Human ovarian tumour protein encoding cDNA #322.
KW Human; ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;
KW gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.
XX Homo sapiens.
XX WO200190154-A2.
XX PD 29-NOV-2001.
XX PF 23-MAY-2001; 2001WO-US016895.
XX PR 24-MAY-2000; 2000US-0207107P.
XX PR 13-JUN-2000; 2000US-0211457P.
XX PR 21-JUN-2000; 2000US-0213673P.
XX PR 03-AUG-2000; 2000US-0223288P.
XX PR 01-MAR-2001; 2001US-0272790P.
XX (CORI-) CORIXA CORP.
XX Xu J, Mitcham JL, Harlocker SL, Dillion DC, Secretist H, Lodes MJ;
XX Algate PA, Fling SP, Mannion J, Benson DR, Carter D;
XX WPI; 2002-097641/13.
XX New isolated polynucleotide encoding polypeptide comprising portion of
PT ovarian tumor protein, useful for detection, diagnosis and therapy of
PT human ovarian cancer.

XX Claim 1; Page 266-268; 285pp; English.
PS
XX The invention relates to an isolated polynucleotide encoding a
CC polypeptide comprising a portion of an ovarian tumour protein. The
CC sequences of the invention are useful for stimulating an immune response
CC and for treating ovarian cancer in a patient. An antigen presenting cell
CC that expresses the sequences is useful for treating ovarian cancer by
CC incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells
CC can then be proliferated and administered to the patient to inhibit the
CC development of cancer. The DNA sequences are useful as probes or primers
CC for nucleic acid hybridisation, to direct expression of a polypeptide in
CC appropriate host cells. Detecting the presence of a cancer in a patient
CC involves obtaining a biological sample from the patient, contacting the
CC biological sample with an agent that binds to the protein, detecting the
CC amount of protein that binds to the agent, comparing the amount of
CC protein to a predetermined cut-off value and determining the presence of
CC cancer. Sequences ABK09464-ABK09802 represent PCR primers and cDNA
CC molecules encoding ovarian tumour proteins of the invention
XX
SQ Sequence 4701 BP; 928 A; 1574 C; 1471 G; 728 T; 0 U; 0 Other;
Query Match 94.3%; Score 3484.6; DB 6; Length 4701;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3490; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 180 GCCTCCCTGCTCCCTGCGAGCTGAGTCCCAACGGGGGCCAGGAGCACGCGGATGAAGAA 239
Db 1203 GCCCGCAAGTACAGAGAGCTGAGTCCCAACGGGGGCCAGGAGCACGCGGATGAAGAA 1262
Qy 240 CGTCCCGGTGCGCGTGTACTGCCGCCCTCTGCTGGAGAAGGACCCACCATGAAGCTGTG 299
Db 1263 CGTCCCGGTGCGCGTGTACTGCCGCCCTCTGCTGGAGAAGGACCCACCATGAAGCTGTG 1322
Qy 300 GTGTGCCCGGGCGCTCAACCTGAGCGGTGAGGCCCAATGAGAGCAGCGTGGGAATGG 359
Db 1323 GTGTGCCCGGGCGCTCAACCTGAGCGGTGAGGCCCAATGAGAGCAGCGTGGGAATGG 1382
Qy 360 AGTCAAGCCAGCGCGCGCGCGATCCCTCAGCTGCGACCGCGAAGGAGACGGCGAGCC 419
Db 1383 AGTCAAGCCAGCGCGCGCGCGATCCCTCAGCTGCGACCGCGAAGGAGACGGCGAGCC 1442
Qy 420 CAAGAGCGCCCAACAGCTCTCCCGAGAGAAGAGCCCAAGGAGCTCCCTGAAATGAGCGC 479
Db 1443 CAAGAGCGCCCAACAGCTCTCCCGAGAGAAGAGAGGCAAGGAGCTCCCTGAAATGAGCGC 1502
Qy 480 CACTCCAGCGCGGTGTGGATCTCGACAGCACCTGTACACCAAGAGTGTGTGATCAT 539
Db 1503 CACTCCAGCGCGGTGTGGATCTCGACAGCACCTGTACCAAGAGTGTGTGATCAT 1562
Qy 540 CGAGCCCAACAGCGCGCGCGCGAGTGGAGCCAGTTTCACTGCTGCAACGCGCAGTGT 599
Db 1563 CGAGCCCAACAGCGCGCGCGCGAGTGGAGCCAGTTTCACTGCTGCAACGCGCAGTGT 1622
Qy 600 GTGCATCTCCAGCATCCCGCGCGCGAGCGACAGCGACTACCTCCCGGGGAGATTTCT 659
Db 1623 GTGCATCTCCAGCATCCCGCGCGCGAGCGACAGCGACTACCTCCCGGGGAGATTTCT 1682
Qy 660 GGAAGCGACGTGAACCCAGAGGACCCCGGGCGCAGATGCGTGTGCGCGGTATCACCT 719
Db 1683 GGAAGCGACGTGAACCCAGAGGACCCCGGGCGCAGATGCGTGTGCGCGGTATCACCT 1742
Qy 720 GGTGGGTGTGCAACCCCGCTGCAAGCGCGGAGCAACTGCTCTCCCGAGGGGACAC 779
Db 1743 GGTGGGTGTGCAACCCCGCTGCAAGCGCGGAGCAACTGCTCTCCCGAGGGGACAC 1802
Qy 780 CCCAGTGTAGACAAAGGGCGAGGGAGGTGGCCACCATGCGCAACGCGGAAGGTCAACCC 839
Db 1803 CCCAGTGTAGACAAAGGGCGAGGGAGGTGGCCACCATGCGCAACGCGGAAGGTCAACCC 1862
Qy 840 GTCCCGTGTGCAAGAGGAGGCGCACAGAGGCGCACCGAGGTGCCAGACCTTGGCCCGAG 899
Db 1863 GTCCCGTGTGCAAGAGGAGGCGCACAGAGGCGCACCGAGGTGCCAGACCTTGGCCCGAG 1922


```

Db 4083 CAGTCTTAGACACAGCTCAGCTCAGCTGAGTGGGGGATGATGCTCCCGGGGCTG 4142
Qy 3120 CCTCTCAGCGGCTCCAGGAGCCAGCTCCAGACAGCTACTAAGTGCCTAGGGTT 3179
Db 4143 CTTCTGACGGGCTCCAGGAGCCAGCTCCAGACAGCTACTAAGTGCCTAGGGTT 4202
Qy 3180 GCGCGTGTGGCTGTCTCCAGGAGGAAACAGAGAGCGCCACCAAGAGAGCCCGTGGG 3239
Db 4203 GCGCGTGTGGCTGTCTCCAGGAGGAAACAGAGAGCGCCACCAAGAGAGCCCGTGGG 4262
Qy 3240 CTGAGGATGGAGCGCGCCCGAGCGACTCCCAAGCCCGAGAGGCGAGCGCCACTGGA 3299
Db 4263 CTGAGGATGGAGCGCGCCCGAGCGACTCCCAAGCCCGAGAGGCGAGCGCCACTGGA 4322
Qy 3300 CTGCTCTCCCTGCGCAGCTGGGCTCTCTGSCCTATTCTTACCTTCCAGGCGCCACTGCAC 3359
Db 4323 CTGCTCTCCCTGCGCAGCTGGGCTCTCTGSCCTATTCTTACCTTCCAGGCGCCACTGCAC 4382
Qy 3360 TCCTGTCTGGAGGCGCTTATGAGGCGAGCCAGCCCGCCCGCACCACCCCAACAGAGA 3419
Db 4383 TCCTGTCTGGAGGCGCTTATGAGGCGAGCCAGCCCGCCCGCACCACCCCAACAGAGA 4442
Qy 3420 AGCACAGATCTGGGAGCTGCCCAAGCCCGCTGCGCAGCGGCTGCGCCGCT 3479
Db 4443 AGCACAGATCTGGGAGCTGCCCAAGCCCGCTGCGCAGCGGCTGCGCCGCT 4502
Qy 3480 GCGTGTCCCGGCTTCTCCACAGCCCTGCCACCTCCTCATGTATGTTCGCTCCCTC 3539
Db 4503 GCGTGTCCCGGCTTCTCCACAGCCCTGCCACCTCCTCATGTATGTTCGCTCCCTC 4562
Qy 3540 GTCTGTCCCGAGATCTCGAAGTACTCGGGCTGAGCAGTGGGGGCTGGGGAGG 3599
Db 4563 GTCTGTCCCGAGATCTCGAAGTACTCGGGCTGAGCAGTGGGGGCTGGGGAGG 4622
Qy 3600 GGTGACGATCTCTCTCAGGCTTGGCCCTGCAAGCAAAACCATATCTGCTCTGTATGT 3659
Db 4623 GGTGACGATCTCTCTCAGGCTTGGCCCTGCAAGCAAAACCATATCTGCTCTGTATGT 4682
Qy 3660 AATAAATGTCTTAAGCTCG 3678
Db 4683 AATAAATGTCTTAAGCTCG 4701

RESULT 5
AAI61066
ID AAI61066 standard; cDNA; 4956 BP.
XX
AC AAI61066;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 5055.
XX
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia; ss.
XX
OS Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2006; 2000WO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.
XX
XX 21-JAN-2000; 2000US-00488725.
XX
XX 25-APR-2000; 2000US-00552317.
XX
XX 20-JUN-2000; 2000US-00598042.

```

```

PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HVSE-) HVSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue A, Yang Y, Zhang J, Zhao QA;
XX Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX P-PSDB; AAM41910.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 5055; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
XX encoded polypeptides (AAM38642-AAM42213) with nontropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, drug screening,
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX Sequence 4956 BP; 1014 A; 1614 C; 1543 G; 785 T; 0 U; 0 Other;
XX
XX Query Match 91.3%; Score 3374; DB 4; Length 4956;
XX Best Local Similarity 98.6%; Pred. No. 0;
XX Matches 3478; Conservative 0; Mismatches 35; Indels 13; Gaps 7;
XX
Qy 180 GCCTCCCTGCTCCCTGTCAGCTGAGTCCCAACGCGGGGCCAGGAGCACGCGGATGAAGAA 239
Db 1194 GCCGCCAAGTACAAAGCAGCTGAGTCCCAACGCGGGGCCAGGAGCACGCGGATGAAGAA 1253
Qy 240 CGTGCCTGCGCGGTGTAAGTCTGCGCCCTCTGCTGAGAGAGCCCAACCATGAAGCTGTG 299
Db 1254 CGTGCCTGCGCGGTGTAAGTCTGCGCCCTCTGCTGAGAGAGCCCAACCATGAAGCTGTG 1313
Qy 300 GTGTGCGCGCGCGTCAACCTGAGCGGTGGAGGCCCAATGAGGACGACGCTGGGAATGG 359
Db 1314 GTGTGCGCGCGCGTCAACCTGAGCGGTGGAGGCCCAATGAGGACGACGCTGGGAATGG 1373
Qy 360 AGTCAAGCCAGCGCCAGCGCGCGATCCCTGTGACCTGCGACCGCGAAGGAGAGCGCGAGCC 419
Db 1374 AGTCAAGCCAGCGCCAGCGCGCGATCCCTGTGACCTGCGACCGCGAAGGAGAGCGCGAGCC 1433
Qy 420 CAAGAGCGCCGACAGCTCTCCCGAGAGAGAGGCGGAGGAGCTCCCTGAATGAGCGC 479
Db 1434 CAAGAGCGCCGACAGCTCTCCCGAGAGAGGCGGAGGAGCTCCCTGAATGAGCGC 1493
Qy 480 CACCTCCAGCGCGGTGTGGATCCTGACACGACCCCTGACACGACGAGGAGGAGGATGATCAT 539
Db 1494 CACCTCCAGCGCGGTGTGGATCCTGACACGACCCCTGACACGACGAGGAGGAGGATGATCAT 1553
Qy 540 CGACGCGCAACAGCGCGGCGACCGGTGTGACCGATTCACCTGCTGCAACGCGCACGCTGCT 599
Db 1554 CGACGCGCAACAGCGCGGCGACCGGTGTGACCGATTCACCTGCTGCAACGCGCACGCTGCT 1613
Qy 600 GTGCATCTCAGCATCCCCCGCGCGCAGCGACGCGACTACCTCCCGGGGAGATGTTCTCT 659
Db 1614 GTGCATCTCAGCATCCCCCGCGCGCAGCGACGCGACTACCTCCCGGGGAGATGTTCTCT 1673

```


QY	478	GCACCTCCAGCC-GGGTGTGGATCTTGACAGACACCTGACACACCAAGGTGG-TGA	535
Db	1506	GCACCTCCAGCCGGGGTGTGGATCTCTGACAGACACCTGACACCAAGGTGGATGA	1565
QY	536	TCATCAGCGCAACACGCGGGCAGCGTGTGTGACACAGTTTCAAGTGTGCAAGCGCAAG	595
Db	1566	TCATCAGCGCAACACGCGGGCAGCGTGTGTGACACAGTTTCAAGTGTGCAAGCGCAAG	1625
QY	596	TGCTGTGCAATCTCAGCATCCCGCGGCAGACGACGCACTACCTCCCGGGGAGATGT	655
Db	1626	TGCTGTGCAATCTCAGCATCCCGCGGCAGACGCACTACCTCCCGGGGAGATGT	1685
QY	656	TCTGTGACAGCGAGTGAACCCAGAGGACCCCGGGCGAGATGCGTGTGCGCCGATCA	715
Db	1686	TCTGTGACAGCGAGTGAACCCAGAGGACCCCGGGCGAGATGCGTGTGCGCCGATCA	1745
QY	716	CCCTGTGTGGCTGTGTCACCCGCTGCAACGTCGCGGAGCACTGCTCTCCCGAGGGG	775
Db	1746	CCCTGTGTGGCTGTGTCACCCGCTGCAACGTCGCGGAGCACTGCTCTCCCGAGGGG	1805
QY	776	ACACCCAGTGTAGCAAGGGGACAGGGGAGTGGCCACCATCGGCCAACGGGAAGGTCA	835
Db	1806	ACACCCAGTGTAGCAAGGGGACAGGGGAGTGGCCACCATCGGCCAACGGGAAGGTCA	1865
QY	836	ACCCGTCCTCAGTCCACAGAGGAGGCCACAGAGGCCACGAGGTGCCAGACCTGGGCCCA	895
Db	1866	ACCCGTCCTCAGTCCACAGAGGAGGCCACAGAGGCCACGAGGTGCCAGACCTGGGCCCA	1925
QY	896	GCGAGCCAGAGACAGCACATTGCGGCCCGCGGCTCTCACAGAGCACGTTTCACTGACC	955
Db	1926	GCGAGCCAGAGACAGCACATTGCGGCCCGCGGCTCTCACAGAGCACGTTTCACTGACC	1985
QY	956	CAGCCCGACCCGCTCTTGCCCGCCAGCCTGGCAGCGAGAACCGGCCAGAGCCTGACA	1015
Db	1986	CAGCCCGACCCGCTCTTGCCCGCCAGCCTGGCAGCGAGAACCGGCCAGAGCCTGACA	2045
QY	1016	GCAGCAGCACAGCCAGAGCCAGAGCCAGCGGGACCCCAAGGAGCAGCAGTGT	1075
Db	2046	GCAGCAGCACAGCCAGAGCCAGAGCCAGCGGGACCCCAAGGAGCAGCAGTGT	2105
QY	1076	CTGACCCACCATGTGGTGGAGCCAGAGCGGCTGCTATGTGCACTGGCTGTGG	1135
Db	2106	CTGACCCACCATGTGGTGGAGCCAGAGCGGCTGCTATGTGCACTGGCTGTGG	2165
QY	1136	CCAACTGGAAGAAGTGCTGCACTCCATCAAGCTGAAGATTTCTGTGTGAGCCTGTGG	1195
Db	2166	CCAACTGGAAGAAGTGCTGCACTCCATCAAGCTGAAGATTTCTGTGTGAGCCTGTGG	2225
QY	1196	ATGTCAAAGGCCGTGTGGTGGCTCTGGCGGAGCGGACCTGGCCATCTTCCACCGTG	1255
Db	2226	ATGTCAAAGGCCGTGTGGTGGCTCTGGCGGAGCGGACCTGGCCATCTTCCACCGTG	2285
QY	1256	GTGAAGATGCGCAGTGGGATCTGAGCAACTATACCTAATGCACTGGGCCACCGCAC	1315
Db	2286	GTGAAGATGCGCAGTGGGATCTGAGCAACTATACCTAATGCACTGGGCCACCGCAC	2345
QY	1316	ACTCCATCCGCTGATGGCTGTGTGACGACCGCGTGTGGTGTGGCTTACAAGAACAGG	1375
Db	2346	ACTCCATCCGCTGATGGCTGTGTGACGACCGCGTGTGGTGTGGCTTACAAGAACAGG	2405
QY	1376	TGCACGTCATCCAGCCCAAGACCATGACAGATAGAGAGTCAATTTGACGCCCAACCGCGG	1435
Db	2406	TGCACGTCATCCAGCCCAAGACCATGACAGATAGAGAGTCAATTTGACGCCCAACCGCGG	2465
QY	1436	GGGAGAGCCAGGTGCGGACGTCGCGTGGATCGCGATGGCGTATGGGTGTCCATCCGCC	1495
Db	2466	GGGAGAGCCAGGTGCGGACGTCGCGTGGATCGCGATGGCGTATGGGTGTCCATCCGCC	2525
QY	1496	TGGAATCCACCTGAGGCTTACCATGACACACGACCAAGCACTTACAGAGAGTGGACA	1555
Db	2526	TGGAATCCACCTGAGGCTTACCATGACACACGACCAAGCACTTACAGAGAGTGGACA	2585
QY	1556	TTGAGCCCTACGTCAGCAAGATGCTAGGCACTGGCAAGCTGGGTTTCTCCTTCGTACGCA	1615

Db	2586	TTGAGCCCTACGTCAGCAAGATCTAGGCACTGGCAAGCTGGGTTTCTCCTTCGTACGCA	2645
QY	1616	TCACGGCCCTGCTGTGTCGCGGCGAGCGGCTCTGGGTGGGCAAGGAGTGTGTCA	1675
Db	2646	TCACGGCCCTGCTGTGTCGCGGCGAGCGGCTCTGGGTGGGCAAGGAGTGTGTCA	2705
QY	1676	TCTCCATCCCTCTGAC-AGAGACTGTGGTCTGTACCCAGGGCCA--GCTCCTCGGGCTCC	1732
Db	2706	TCTCCATCCCTCTGACAGAGACTGTGGTCTGTACCCAGGGCCAAGCTCTCGGGCTCC	2765
QY	1733	GAGCCAAATAGACATCCCCACCTCTGGGG-AGGGCGCCGCTCCCGGGGCG-ATCATCCA	1790
Db	2766	GAGCCAAATAGACATCCCCACCTCTGGGGAAAGGGGCGCCGCTCCCGGGGCGAATCATCCA	2825
QY	1791	CGTGTAT-GGCGATGACAGCAGTGAACAGGGCGGCGAGAGCTTTCATCCCTACTGCTCCA	1849
Db	2826	CGTGTATGGCGATGACAGCAGTGAACAGGGCGGCGAGAGCTTTCATCCCTACTGCTCCA	2885
QY	1850	TGGCCAGGCCCTCAGCTATGCTTCCATGGGCACCGCGATCCGCTGAAGTTTCTTGTCTCG	1909
Db	2886	TGGCCAGGCCCTCAGCTATGCTTCCATGGGCACCGCGATCCGCTGAAGTTTCTTGTCTCG	2945
QY	1910	TGCCAGGGGAACGTGTGTCGCCACCCCTGAATGGCAGTGTGTGACAGCCAGCGAGGCC	1969
Db	2946	TGCCAGGGGAACGTGTGTCGCCACCCCTGAATGGCAGTGTGTGACAGCCAGCGAGGCC	3005
QY	1970	CTGGGCCAGCTGCCCTCGTGGAGGTGAGGGGCCAGAGCTGCGGACGCTGTGGTGC	2029
Db	3006	CTGGGCCAGCTGCCCTCGTGGAGGTGAGGGGCCAGAGCTGCGGACGCTGTGGTGC	3065
QY	2030	TGAGCGCGGGGAGGCTACATCGACTTTCGATTTGGAGACGAGAGGACGACGAGACGG	2089
Db	3066	TGAGCGCGGGGAGGCTACATCGACTTTCGATTTGGAGACGAGAGGACGACGAGACGG	3125
QY	2090	AGGAGGGCGAGGAGACATGACGAGTGAAGCCCGTGTCTCAAGGAGAGCCAGTTC	2149
Db	3126	AGGAGGGCGAGGAGACATGACGAGTGAAGCCCGTGTCTCAAGGAGAGCCAGTTC	3185
QY	2150	ACATCATCGTGTGGCAGGTGCTTACACCCCGAGTGAAGCTGCTGCCCTGCTGGCCCG	2209
Db	3186	ACATCATCGTGTGGCAGGTGCTTACACCCCGAGTGAAGCTGCTGCCCTGCTGGCCCG	3245
QY	2210	ACCTGTACATGAGACCCCGACCACTGACCCCGCCCGCGGCGGTAGCCAGCCAG	2269
Db	3246	ACCTGTACATGAGACCCCGACCACTGACCCCGCCCGCGGCGGTAGCCAGCCAG	3305
QY	2270	GCGCGCGCGCCCTTCTTACCTCTCAACCTGAGCTTTTACCTGAGTGTGGCCCTCC	2329
Db	3306	GCGCGCGCGCCCTTCTTACCTCTCAACCTGAGCTTTTACCTGAGTGTGGCCCTCC	3365
QY	2330	AGCGGGCAGGAGTGGCGGATGCGGATCAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGG	2389
Db	3366	AGCGGGCAGGAGTGGCGGATGCGGATCAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGG	3425
QY	2390	CCGAGGGGAGAGTGTCTCTCGGACAGTTTCCCGGCGAGTCTCTGGCCAGCTTCCAGGCC	2449
Db	3426	CCGAGGGGAGAGTGTCTCTCGGACAGTTTCCCGGCGAGTCTCTGGCCAGCTTCCAGGCC	3485
QY	2450	AGAGTCTCTCAAGTCCAGGGCACCTTGGGCCAGCGGACGAGGAGGAGGAGGAGGAGGAGG	2509
Db	3486	AGAGTCTCTCAAGTCCAGGGCACCTTGGGCCAGCGGACGAGGAGGAGGAGGAGGAGGAGG	3545
QY	2510	CTTACCTCTGGGCTCTTACTTCCCGACACCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2569
Db	3546	CTTACCTCTGGGCTCTTACTTCCCGACACCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3605
QY	2570	GAGGCTGCTGCTGGGCCCACTCTGATGCTGTCTATGGGGCCACCTGCTGCTCTG	2629
Db	3606	GAGGCTGCTGCTGGGCCCACTCTGATGCTGTCTATGGGGCCACCTGCTGCTCTG	3665
QY	2630	GCCTCAGCTGCTAGGGAGGCTGGGCCAGGCACTAGCCTTTGCCAGGGAGGAGGAGGAGGAGG	2689

Db	3666	GCCTCACTCTGCTTAGGGAGCTGGGCCCAGGCACCTAGCCTTTGCCCAAGGAGGTGGGCC	3725
Qy	2690	TCAGGTCGCCAGGTGCTGCACCCACAGCGGCCCTTCTGTGGGCTCCCCGTGTCAG	2749
Db	3726	TCAGGCTGGCCAGGTGCTGTCACCCAGCGCGCTTCTGTGGGCTTCCCCTGTCAG	3785
Qy	2750	CCCTATCCCTGTCTGTCCCCACCCACAGCTGTCCCTTGCCCCAGGAGCTGGCATAAAGCA	2809
Db	3786	CCCTATCTCTGTCTGTCCCCACCCACAGCTGTCCCTTGCCCCAGGAGCTGGCATAAAGCA	3845
Qy	2810	CGAGGCCCGCTCCCTGTGGGCGAGCTCTTGAGAAACAGAGACTGCTACCCCATCTCG---CCC	2868
Db	3846	CGAGGCCCGCTCCCTGTGGGCGAGCTCTTGAGAAACAGAGACTGCTACCCCATCTCGAGCCC	3905
Qy	2869	ATGCAGCAGCAGCTCTTGCCAGACCCCGTTCTGACCCGTGTCCCGCCAGGCTCTGCCTGGGC	2928
Db	3906	ATGCAGCAGCAGCTCTTGCCAGACCCCGTTCTGACCCGTGTCCCGCCAGGCTCTGCCTGGGC	3965
Qy	2929	AGAAAGCTACCTTGGAGGAGTGGGCCCTGGAGTCTTGTCCCTCCAGAAAGCCCCCAGGG	2988
Db	3966	AGAAAGCTACCTTGGAGGAGTGGGCCCTGGAGTCTTGTCCCTCCAGAAAGCCCCCAGGG	4025
Qy	2989	TGSGATTTCTCAGGCTGCCAGGCGAGGCCCAGAGCCTCAGAAAGGAGGAGGCCCTGGC	3048
Db	4026	TGGATTTCTCAGGCTGCCAGGCGAGGCCCAGAGCCTCAGAAAGGAGGAGGCCCTGGC	4085
Qy	3049	CTCTCCGGGATCAGTCTCTAGGACACAGGCTCAGCCTCAGCTTGATGGGGATGATGTCT	3108
Db	4086	CTCTCCGGGATCAGTCTCTAGGACACAGGCTCAGCCTCAGCTTGATGGGGATGATGTCT	4145
Qy	3109	CCCGGGCCTGCTCTGTCACGGGGCTCCAGGAGCCACAGTCCACAGACAGCTACTAAG	3168
Db	4146	CCCGGGCCTGCTCTGTCACGGGGCTCCAGGAGCCACAGTCCACAGACAGCTACTAAG	4205
Qy	3169	TGCTTAGGCTTCCCGCTGTGGCTGTCTGCCAGGAGCAACAGAGAGGCCAACAGCAGA	3228
Db	4206	TGCTTAGGCTTCCCGCTGTGGCTGTCTGCCAGGAGCAACAGAGAGGCCAACAGCTGA	4265
Qy	3229	GGCCCGTGGGCTGAGGATGAGCGCCGCCAGCCGACTCCAGCCCGCAGAGGGCAGAC	3288
Db	4266	TGCCCGTGGGCTGAGGATGAGCGCCGCCAGCCGACTCCAGCCCGCAGAGGGCAGAC	4325
Qy	3289	GCACCCCTGGACTCTCTCCCTGCCAGCTGGGCTCTCTGGGC---TATTTCTTACCTT	3344
Db	4326	GCACCCCTGGACTCTCTCCCTGCCAGCTGGGCTCTCTGGGCCTTCTTCTTACCTT	4385
Qy	3345	CCAGGCCCACTGC-ACTCCTGTCTGGGAGGCCCTTATGAGGGCAGCCC---AGCCCCGC	3400
Db	4386	CCAGGCCCACTGCAACTCTCTGTCTGGGAGGCCCTTATGAGGGCAGCCCGCCCCCGG	4445
Qy	3401	ACGCACCCCAACACGAGNAGCAGACTTTGGGAGCTGCCCCACAAAG---CCCGCTGGC	3459
Db	4446	CACCAACCCCAACACGAGNAGCAGACTTTGGGAGCTGCCCCACAAAGCCCGCTGGC	4505
Qy	3460	CACCGA---GGGCTGCAGCCGCTCGCCTGCC---GGCTTCTCCCCACCAACCTGCACCTCC	3515
Db	4506	CACCGAAGGCTTGCAATGCGTGCCTGCCCGCTTCTCCCGCACCACCTTGCCACCTCC	4565
Qy	3516	ACTGTGATGATGTCCGCTCCCTGCTGTTCCTCCCCCAGAGTCTCGAGTGCATCCGGCT	3575
Db	4566	ACTGTGATGATGTCCGCTCCCTGCTGTTCCTCCCCCAGAGTCTCGAGTGCATCCGGCT	4625
Qy	3576	GACGAGTGGGCGGCTGGGCGAGGGGTACGATTCTCCTCAGGCTTTGGCCCTGCAAGCA	3635
Db	4626	GACGAGTGGGCGGCTGGGCGAGGGGTGACGATTCTCCTCAGGCTTTGGCCCTGCAAGCA	4685
Qy	3636	AAACCCACATATCTCTCTGTATGTATAAATGTCTTTAAGCTCG	3678
Db	4686	AAACCCACATATCTCTCTGTATGTATAAATGTCTTTAAGCTCG	4728

AA62899	standard; DNA; 4173 BP.
AA62899	
AA62899	
01-NOV-2000	(first entry)
Murine JNK3 binding protein DNA sequence.	
JNK3; Jun N-terminal protein kinase 3; JNK3 binding protein; JSAP; JNK/SAPK associated protein; mitogenic; neuroprotective; neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective; antiinflammatory; nervous system disease; Alzheimer's disease; Parkinson's disease; Huntington's disease; multiple sclerosis; muscular dystrophy; apoplexy; schizophrenia; dropsy; inflammatory disorders; mouse; ds.	
Mus sp.	
Key	Location/Qualifiers
CDS	107..4024
FT	/*tag= a
FT	/product= "JSAP"
FT	/note= "JNK3 binding protein"
W0200031132-A1.	
02-JUN-2000.	
19-NOV-1999;	99WO-JF006487.
24-NOV-1998;	98JP-00332484.
02-SEP-1999;	99JP-00248442.
(KYOW) KYOWA HAKKO KOGYO KK.	
Ichimura M, Hirose R, Yoshioka K;	
WPI; 2000-400042/34.	
P-PSDB; AAB12875.	
Polypeptides binding to Jun N-terminal protein kinase for treatment and diagnosis of nervous system and inflammatory disorders.	
Disclosure; Page 110-123; 254pp; Japanese.	
The invention relates to polypeptides which bind to JNK3 (Jun N-terminal protein kinase 3). The present sequence represents a DNA sequence encoding a JNK3 binding protein (JSAP; JNK/SAPK associated protein). Included in the invention are polypeptides derived from the JNK3 binding proteins, DNA sequences encoding the proteins, expression vectors containing the DNA, and host cells transformed by the vectors. The polypeptides exhibit mitogenic, neuroprotective, antiparkinsonian, anticonvulsant, cerebroprotective, neuroprotective, and antiinflammatory activity. JNK3 binding proteins are used in the treatment, prevention and diagnosis of diseases associated with the JNK3 cascade, such as diseases of the nervous system including Alzheimer's disease, Parkinson's disease, Huntington's disease, multiple sclerosis, muscular dystrophy, apoplexy and schizophrenia, dropsy, and inflammatory disorders	
Sequence 4173 BP; 1027 A; 1146 C; 1214 G; 786 T; 0 U; 0 Other;	
Query Match	41.5%; Score 1531.6; DB 3; Length 4173;
Best Local Similarity	82.3%; Pred. No. 2.9e-289;
Matches 1785; Conservative	0; Mismatches 379; Indels 6; Gaps 2;
148	GATCGGTAGGAGCCAGGGTTGTCGCCACGGGGCTCCCTGCTCCTCGAGTGAATGCC 207
1985	GACGGAGGCTGTCAGGCGCTGTGGGTGGAGCTGCTGCCAAGTCAAGCAGCTGAGCCCC 2044
208	AACGGGGCCAGGAGGACACCGGATGAAGAACGTGCCGGTGCCGGTGTAATGCCGCCCT 267
2045	AATGAGGCCAGGAGACACCGGATGAAGAACGTGCCGGTGTAATGCCCTGTGTACTGTGCGCCT 2104
268	CTGTGGAGAGGACCCCAACCATGAGCTGTGGTGTGGCGGGCGGTCAACCTGAGCGGG 327

DT	01-NOV-2000	(first entry)	Db	2192	TGGAAGCCATGAGAGGACTCTAGCAATGGACCAAGCCTGTACCCAGGTCCGAGACCCCT	2251
XX						
DE						
XX						
KW	Murine JNK3 binding protein DNA sequence #2.		Qy	388	CTGACCTGCGACCGCGAGAGAGAGCGGAGCCCAAGAGCGCCACACAGTCTCCCGAGAAG	447
KW	JNK3; Jun N-terminal protein kinase 3; JNK3 binding protein; JSAP;		Db	2252	CTGACCTGTGACCGGGAAGAGCGAAACCCCAAGAGACACACCCATCACCCTGAGAGAG	2311
KW	JNK/SAPK associated protein; neurotropic; neuroprotective; neuroleptic;		Qy	448	AAGAGGCCCAAGAGAGCTCCCTGAAATGGAGCGACCACTCCAGCGGGGTGTGATCCTGACC	507
KW	antiparkinsonian; anticonvulsant; cerebroprotective; antiinflammatory;		Db	2312	AAGAAGCAAGAGAAACCCCTGAGGAGATGCTACCTCCAGTCGGGTATGGATCCTCACC	2371
KW	nervous system disease; Alzheimer's disease; Parkinson's disease;		Qy	508	AGCACCTTGACCCACCAAGAGTGGTGTATCATCGACCGCCAAACAGCCGGGACCGTGGTG	567
KW	Huntington's disease; multiple sclerosis; muscular dystrophy; apoplexy;		Db	2372	AGCACCTTGACACCAAGAGTGGTGTATCATTCATGCAACCAAGCGAGGACCAATTTGTG	2431
XX	schizophrenia; dropsy; inflammatory disorders; mouse; ds.		Qy	568	GACCACTTACCGCTGTGCAACCGGACAGTGTGTGCACTCCAGCACTCCCGCGGCCAGC	627
OS	Mus sp.		Db	2432	GATCAGTTTCAAGTCTGCAATGCCCCACGTCCTGTGTATCTCCAGCAATTCCTGCGGCCAGT	2491
XX						
XX	Key	Location/Qualifiers	Qy	628	GACAGCGACTACCTCCCGGGAGATGTTCTGACACAGCGACGTCGAACCCAGAGAGACCG	687
XX	CDS	107..4052	Db	2492	GACAGTGAATATCCCTCGGGAGATGTTCTAGACAGTGTGTAACCTGAAGATTCA	2551
FT	/*tag= a		Qy	688	GGCGCAGATGGCGTGTGCGCGGTATCACTCTGGTGGGTGTGCCACCCGCTGCAACGTG	747
FT	/product= "JSAP"		Db	2552	GGTCTGATGGTGTGCTGGCTGGCATCACCTCGTGGGTGTGCTTACCCGCTGCAATGTT	2611
FT	/note= "JNK3 binding protein"		Qy	748	CCGCGGAGCAACTGCTCTCCCGAGGGGACACCCAGTGTGTAGACAAAGGGGACAGGGGAG	807
XX	WO200031132-A1.		Db	2612	CCAGGTAGCAACTGTTCCTACGAGGAGACACCCAGTACTGGACAAGGGGACGGGGAT	2671
XX	02-JUN-2000.		Qy	808	GTGGCCACCATCGCAACGGGAAGTCAACCCGTCCTCCAGTCCACAGAGAGAGCCACAGAG	867
XX	19-NOV-1999; 99WO-JP006487.		Db	2672	GTGGGACCATCTGCCAATGGGAAGGTCAACCGTCCCAATCCACAGAGAAGAGCCACAGAA	2731
XX	24-NOV-1998; 98JP-00332484.		Qy	868	GCCACGAGGTGCCAGACCTCTGGGCCACAGCGAGCCAGAGACAGCACATTTGGGGCCGGG	927
XX	02-SEP-1999; 99JP-00248442.		Db	2732	GCCACAGAGTGGCAACGCTGCTCCAGCGAGTCAGAAGCAACACAGTCCGGGCCCGGG	2791
XX	(KYOW) KYOWA HAKKO KOGYO KK.		Qy	928	CCTCTCAGAGCAGCTTCTCACTGACCCAGCCGACCCGCTCTCTGGCCCCCAGCCT	987
XX	Ichimura M, Hirose R, Yoshioka K;		Db	2792	CCTCTCAGAGCAGCTTCTTACTGACCCAGCCGACCCGCTCTCTGGCCCCCAGCCT	2851
XX	WPI; 2000-400042/34.		Qy	988	GSCACGAGAAACCGGCCAGAGCCTTCACAGCAGCAGCAGCACGCGCCAGAGCCAGCAGC	1047
XX	P-PSDB; AAB12876.		Db	2852	GCCAGTGAGATGGTTCAGATCCATGACCACTTGTACAGCTCAGGTGGAGCCAGT	2911
XX	Polypeptides binding to Jun N-terminal protein kinase for treatment and		Qy	1048	GGGGACCCACCGGAGCAGGAGAGTGTGACCCACCATGTGACCCACCATGTGCTGGAGCCCA	1107
XX	diagnosis of nervous system and inflammatory disorders.		Db	2912	GGGGAACTCTCAACAACCAACCAAGTAGCGCTGCACCCCACTATGTGGCTAGGAGCCCA	2971
XX	Claim 4; Page 123-136; 254pp; Japanese.		Qy	1108	GGCTGGCTTATGTGCTACCTCGGCTGTGGCAACTGGAAGAGTCTGCTGCACTCCATCAAG	1167
XX	The invention relates to polypeptides which bind to JNK3 (Jun N-terminal		Db	2972	GGCTGGCTTATGTGCTATTCAGCGGTAGCCAACTGGAAGAGTGTGCTGCACTCCATCAAG	3031
XX	protein kinase 3). The present sequence represents a DNA sequence		Qy	1168	CTGAAGGATTTCTGTGCTGAGCCTGTGTGATGTCAAGGGCCGCTGTGCTGTGGCTTGGCG	1227
XX	encoding a JNK3 binding protein (JSAP; JNK/SAPK associated protein).		Db	3032	CTGAAGACTCTGTGCTGAGCCTGGTGCATGTCAAGGGCCGAGTGTGCTGTGCTTTGCA	3091
XX	Included in the invention are polypeptides derived from the JNK3 binding		Qy	1228	GACGGGACCTCGGCATCTTCCACCGTGTGAGATGGCCAGTGGGATCTTGAGCACTAT	1287
XX	proteins, DNA sequences encoding the proteins, expression vectors		Db	3092	GATGGGACCTCGGCTATCTTCCATCGTGGAGAGGATGGCCAGTGGGACCTTGAGCAACTAC	3151
XX	containing the DNA, and host cells transformed by the vectors. The		Qy	1288	CACTTAATGGACTGGGCCACCGCACCTCCATCCGCTGATGGCTGTGTGTGATGAC	1347
XX	polypeptides exhibit neurotropic, neuroprotective, antiparkinsonian,		Db	3152	CACCTAATGGACTGGGCCACCGCACCTCCATCCGCTGATGGCTGTGTGTGATGAC	3211
XX	anticonvulsant, cerebroprotective, neuroleptic, and antiinflammatory		Qy	1348	CGCGTGTGTGTGGCTTACAGAACAAAGGTGCAGCTCATCCAGCCCAAGACCATGACATA	1407
XX	activity. JNK3 binding proteins are used in the treatment, prevention and		Db	3212	CGAGTTTGTGTGGCTTACAGAACAAAGGTGCATGTTATCCAGCCCAAGACCAATGACAT	3271
XX	diagnosis of diseases associated with the JNK3 cascade, such as diseases		Qy	1408	GAGAACTTATGACGCCCAACCGCGCGGGAGAGCCAGTGGCGGAGCTGGCGTGGATC	1467
XX	of the nervous system including Alzheimer's disease, Parkinson's disease,					
XX	Huntington's disease, multiple sclerosis, muscular dystrophy, apoplexy					
XX	and schizophrenia, dropsy, and inflammatory disorders					
XX						
XX	Sequence 4200 BP; 1034 A; 1153 C; 1223 G; 790 T; 0 U; 0 Other;					
XX						
XX	Query Match	41.5%; Score 1531.6; DB 3; Length 4200;				
XX	Best Local Similarity	82.3%; Pred. No. 2.9e-289;				
XX	Matches 1785; Conservative	0; Mismatches 379; Indels				
XX		6; Gaps				
XX		2;				
XX						
XX	148	GATGGGTAGAGCAGGGTTCGTGCCCAAGAGTGGAGCCTCCCTGCTCCCTGAGCTGAGTCCC	207			
XX	2012	GACGGGAGGCTGAGGCGCTGTGGGTGGAGCCTCCCTGCGCAAGTACAGAGCTGAGCCCC	2071			
XX	208	AACGGGGGCGCAGGAGACACCGGATGAAGAACTGCCGGTGGCGGTGTACTGCCGCCCT	267			
XX	2072	AATGGAGCCAGAGAGACACCGGATGAAGAAATGTGCTGTCTGTTACTGTGCCCT	2131			
XX	268	CTGGTGAGAGAGACCCACCATGAAGCTGTGGTGGCGCGGGGTCAACCTGAGGGG	327			
XX	2132	CTGGTGAGAGAGACCCCTTCGACAAAGCTGTGGTGTGCTGCTGTGTGTCACACCTGAGTGG	2191			
XX	328	TGGAGGCCCAATGAGGACGAGCTGGGAATGGAGTCAAGCCAGCGCCGCGGATCCC	387			

Db	3272	GAGAAATCAATTTGATGCCACCAAGAGGGGGAAAGCCAGGTACGTCAGCTGGCTGGATC	3333
QY	1468	GGCGATGGCGGTATGGGTGTCCATCCGCTGGACTCCACCCCTGAGGCTTACCATGTCAAC	1527
Db	3332	GGTGAATGAGATGTGGGTCTCTAATTCGTTGGATTCTACCCCTTCGGCTCTACCATGCTCAC	3391
QY	1528	ACGCACCAAGCATCTACAGGACGTGGACATTTGAGCCCTTACGTCAGCAAGATGCTAGGCAC	1587
Db	3392	ACCCACCAAGCACTGCAGGATGTGACATGAGCCCTATGTTAGCAAGATGCTAGGAAC	3451
QY	1588	GGCAAGCTGGGTTTCTCTCTTCCGTAACGATCAAGCGCCCTGCTTGTGCGGGCAGCCGGCTC	1647
Db	3452	GGCAAGCTGGGCTTCTCTCTTCTGCGCATACAGGCTTACTCATTTGCAGGCAACCGTCTG	3511
QY	1648	TGGGTGGGCACCGGCAACGGAGTGTCTATCTCCATCCCTCTGACAGAGACTGTGTGCTCTG	1707
Db	3512	TGGGTGGGCACCTGGCAATGGGGTGTGTATCTCCATCCCTCTGACTGAGACTGTGTGCTCTG	3571
QY	1708	CACCAAGCCAGCTCTCTGGGGCTCCGAGCCCAATAAGACATCCCCCACTCTGCGGGAGGGC	1767
Db	3572	CATCAGGCCAGCTCTTAGGGCTCCGAGCCAAAGACATCCCCCAATCTGCGGGAGGGG	3631
QY	1768	GCCCGTCCGGGGGATCATCAAGTGTATGGGATGACAGCATGACAGGGGGGGCAGC	1827
Db	3632	ACCCCGCCAGGGGGCATCATCATGTGTATGGGACGACAGCAGTGACAGGCCCGCAGT	3691
QY	1828	AGCTTCATCCCTACTGCTCTCCATGGCCAGCCAGCTATGCTTCATGGGCAACCGCAT	1887
Db	3692	AGTTTCATCCCTTACTGCTCATGGCACAGGCTCAGCTTTGCTTCCATGGGCACCGTGAT	3751
QY	1888	GCCGTGAAGTTCTTTGTCTCGGTGCCAGGGAACTGCTGTGSCCACTCTGATGCACTGTG	1947
Db	3752	GCTGTCAAATTTCTTTGTCTGTGTGCAGGAATGTGCTGGCCACTCTCAATGGCAGTGTG	3811
QY	1948	CTGACACCCAGCCGAGGGCCCTGGGCCAGCTGCCCTGCTCGAGGTTCGAGGGCCAG	2007
Db	3812	CTAGACAGCCCATCAGAGGGCCCTGGGCCCTGCTGCACCCGCTGCAGATGCTGAGGGCCAG	3871
QY	2008	AAGCTGGGAACTGCTGTGTGTGCTGAGCGCGGGAGGGCTACATCGACTTCCGCATTTGA	2067
Db	3872	AAGTTGAAGAAATGCACTGGTGTGCTGAGTGTGTGAAGGTTACATTGACTTCCGTATCGGA	3931
QY	2068	GACGGAGACACGACGAGACGAGAGGGCGCAGGGGACATGAGCCAGTGAAGCCCGTG	2127
Db	3932	GACGGAGAGATGATGAACTGAGGAATGTGCCGGGACGTGAACACAGAAAGCCCTCG	3991
QY	2128	CTGTCCAAAGCAGAGCCAGTCAATCATCGTGTGGCAGGTGTCTTACACCCCGAGTGA	2187
Db	3992	TTGTTCGAAGCTGAGCGCAGCCACATCATCGTGTGCGAGGTGTCTTACACCCCTGAGTG	4050
QY	2188	AGCTGTGCTCCCTGGTGGCCGACCTGTATAGGACCCCGACACCTGACCCCCCGGCC	2247
Db	4051	AGACCTGTCTCTACCTGATGCCAACTGTATAGGACCCCTTCTGCTGCTCTCCCGCT	4110
QY	2248	GGCCCGCGGGGTAGCCAGCCAGCGCGCGCCCTCTTCTTAACCTCTCAACCTGACGT	2307
Db	4111	GTTCCTGGGGCAGCCA-----GTTTGTGTCATCCCTTTTAACTCTCAACTTCAGCT	4165
QY	2308	TTCACTGAG	2317
Db	4166	TTTGCCTGAG	4175
RESULT 9			
AAA62902			
ID	AAA62902 standard; DNA; 4266 BP.		
XX			
AC	AAA62902;		
XX			
XX	01-NOV-2000 (first entry)		
DE	Murine JNK3 binding protein DNA sequence #4.		
XX			

KW	JNK3; Jun N-terminal protein kinase 3; JNK3 binding protein; JSAP;
KW	JNK/SAPK associated protein; nootropic; neuroprotective; neuropleptic;
KW	antiparkinsonian; anticonvulsant; cerebroprotective; antiinflammatory;
KW	nervous system disease; Alzheimer's disease; Parkinson's disease;
KW	Huntington's disease; multiple sclerosis; muscular dystrophy; apoplexy;
KW	schizophrenia; dropsy; inflammatory disorders; mouse; ds.
OS	Mus sp.
XX	
FH	Key
FT	Location/Qualifiers
FT	107.4117
FT	/tag= a
FT	/product= "JSAP"
FT	/note= "JNK3 binding protein"
XX	
XX	WO200031132-A1.
PN	
XX	
PD	02-JUN-2000.
XX	
PF	19-NOV-1999; 99WO-JP006487.
XX	
PR	24-NOV-1998; 98JP-00332484.
PR	02-SEP-1999; 99JP-00248442.
XX	
PA	(KYOW) KYOWA HAKKO KOGYO KK.
XX	
PI	Ichimura M, Hirose R, Yoshioka K;
XX	
DR	WPI; 2000-400042/34.
DR	P-PSDB; AAB12878.
XX	
PT	Polypeptides binding to Jun N-terminal protein kinase for treatment and
PT	diagnosis of nervous system and inflammatory disorders.
XX	
PS	Claim 4; Page 150-164; 254pp; Japanese.
XX	
CC	The invention relates to polypeptides which bind to JNK3 (Jun N-terminal
CC	protein kinase 3). The present sequence represents a DNA sequence
CC	encoding a JNK3 binding protein (JSAP; JNK/SAPK associated protein).
CC	Included in the invention are polypeptides derived from the JNK3 binding
CC	proteins, DNA sequences encoding the proteins, expression vectors
CC	containing the DNA, and host cells transformed by the vectors. The
CC	polypeptides exhibit nootropic, neuroprotective, antiparkinsonian,
CC	anticonvulsant, cerebroprotective, neuropleptic, and antiinflammatory
CC	activity. JNK3 binding proteins are used in the treatment, prevention and
CC	diagnosis of diseases associated with the JNK3 cascade, such as diseases
CC	of the nervous system including Alzheimer's disease, Parkinson's disease,
CC	Huntington's disease, multiple sclerosis, muscular dystrophy, apoplexy
CC	and schizophrenia, dropsy, and inflammatory disorders
XX	
SQ	Sequence 4266 BP; 1045 A; 1176 C; 1245 G; 800 T; 0 U; 0 Other;

	Query Match	41.5%;	Score 1531.6;	DB 3;	Length 4266;
	Best Local Similarity	82.3%;	Pred. No. 2.9e-289;		
	Matches 1785;	Conservative	0;	Mismatches 379;	Indels 6; Gaps 2;
Qy	148	GATGGGTAGAGCGCAGGGTTCTGTSCCCACGCGCCCTCCCTGCTCCCTGCACTGAGTGCC	207		
Db	2078	GACGGAGGCTGCAGGCTCTGGGTGGAGCCTGCCTGCCAAGTACAAGCACTGAGCCCC	2137		
Qy	208	AACGGGGCCAGAGAGACACCGGATGAGAACTGCGCGGTGCGGTGTACTGCCGCCCT	267		
Db	2138	AATGGAGGCCAGAAAGACACCCGGATGAAAATGTGCCTGTCCCTGTACTGTGCGCCCT	2197		
Qy	268	CTGGTGGAGAGGACCCCACTGAAGCTGTGGTGTGCCGCGGGCGCTCACTTGACGGGG	327		
Db	2198	CTGGTGGAGAGGACCCTTCGACAAAGCTGTGGTGTGCTGTGGTGTCAACTGAGTGGG	2257		
Qy	328	TGGAGGCCCAATGAGGACGACGCTGGGAATGGAGTCAAGCCACGCGCCAGCGCGGATCCC	387		
Db	2258	TGGAAGCCACATGAAGAGACTCTAGCAATGAGACCCAGCCCTGTTCAGGTGAGACCCCT	2317		
Qy	388	CTGACCTGGACCGCGAAGGAGACGGCGAGCCCAAGAGCGGCCACACAGCTCTCCCGAAG	447		

3461	ACCACACGACCTGCGAGGATGTGACATTTGAGCCCTATGTGTAGCAAGATGCTAGGAACC	3520	XX
1588	GGCAAGCTGGGTTCTCTTCCTGATGAGCATACGGCCCTGCTGTTCGCGGCGAGCCGGCTC	1647	XX
3521	GGCAAGCTGGGTTCTCTTCCTGATGAGCATACGGCCCTGCTGTTCGCGGCGAGCCGGCTC	3580	PR
1648	TGGGTGGGACCGGCAACGGAGTGGTTCATCTCCATCCCTCCACAGAGCTGTGTCTCTG	1707	QY
3581	TGGGTGGGACTGGCAATGGGTTGTTCATCTCCATCCCTCCACAGAGCTGTGTCTCTG	3640	Db
1708	CACCGAGGCGAGTCTCTGGGCTCCAGCAATAGACATCCCTCCACAGAGCTGTGTCTCTG	1767	QY
3641	CATCGAGGCGAGTCTCTGGGCTCCAGCAATAGACATCCCTCCACAGAGCTGTGTCTCTG	3700	Db
1768	GCCGCTCCGGGGCATCTCCAGTGTATGGGATGACAGAGTGTGACAGGCGGCGCAGC	1827	QY
3701	ACCCTGGGCGAGGCGCATCTCCATGTGTATGGGACGACAGAGTGTGACAGGCGGCGCAGT	3760	Db
1828	AGCTTCATCCCTACTGCTCCATGGGCGGCGGCGGCTATGCTTCCATGGGCAACCGCAT	1887	QY
3761	AGTTTCATCCCTACTGCTCCATGGGCGGCGGCTATGCTTCCATGGGCAACCGCAT	3820	Db
1888	GCGGTGAAGTCTCTTGTCTGGTGCAGGAACTGTCTGGCCACCTGATGATGCACTGNG	1947	QY
3821	GCTGTCAAAATCTTGTCTGTGCGGAGGAAATGTGTGGCCACTCTCAATGGGCACTGTG	3980	Db
1948	CTGACAGCCGAGCGAGGCGCTGGGCGAGTGGCCCTGCTCGGAGGTTCAGGCGCCAG	2007	QY
3881	CTGACAGCCGAGCGAGGCGCTGGGCGAGTGGCCCTGCTCGGAGGTTCAGGCGCCAG	3940	Db
2008	AAGCTGCGGAACGTGCTGCTGAGCGGCGGAGGGCTTACATCGACTTCCGCATTTGA	2067	QY
3941	AAGTTGAAGATGCACTGGTCTGAGTGGTGGTGAAGTTTACATTTGACTTCCGTATCGGA	4000	Db
2068	GACGAGAGGACGACGAGACGAGGAGGCGCGAGGACATGAGCCAGGTGAGACCCGCTG	2127	QY
4001	GACGAGAGGATGATGAATGAGGAATGTGCGGGGACGTGAACCCAGACAAAGCCCTCG	4060	Db
2128	CTGTCCAGGCGAGCGGAGTACATCATCTGTGGCAGGTGCTTACACCCCGGAGTGA	2187	QY
4061	TTGTCCAGGCTGAGCGGAGCCATCATCTGTGGCAGGTGCTTACACCCCGGAGTGA	4119	Db
2188	AGCTGCTGCCCTGCTGCGGCGGAGCTGTACATAGGACCCCGGACCCCTGACCCCGGCC	2247	QY
4120	AGACCTGTCTTACTGATGATGCACTGTACATAGGACCCCTGACCCCGGCCCT	4179	Db
2248	GGCCCGGGGTAGCAGCGAGGCGCGCGCCCTCTTCACTCTCAACCTGACGCT	2307	QY
4180	GTTCCTGGGCGAGCA-----GGTTGCTCATCCCTTTTAACTCTCAACTTGCAGCT	4234	Db
2308	TTCACTGAG	2317	QY
4235	TTTGCTGAG	4244	Db
RESULT 11			
AAS74089			
ID	AAS74089	standard; cDNA; 2431 BP.	QY
XX	AAS74089;		
AC	AAS74089;		
DT	13-FEB-2002 (first entry)		
XX	DNA encoding novel human diagnostic protein #9893.		
DE	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.		
XX	Homo sapiens.		
OS	WO200175067-A2.		
PN	11-OCT-2001.		
PD			

Db 1103 CCCCCCTCTCTGCCCCCAGCTTGGAGCGAGAACGGGCGAGAGCTTGACAGCAGCAGCA 1162
QY 1025 CACGGCCAGAGCCAGAGCCCGAGCGGGGACCCACCGGAGCAGCAGCAGTGTGACCCCA 1084
Db 1163 CACGGCCAGAGCCAGAGCCCGAGCGGGGACCCACCGGAGCAGCAGCAGTGTGACCCCA 1222
QY 1085 CCATGTGGTGGAGCCAGAACCGGCTGGCTTATGTGCACTCGGCTGTGGCCAACTGGA 1144
Db 1223 CCATGTGGTGGAGCCAGAACCGGCTGGCTTATGTGCACTCGGCTGTGGCCAACTGGA 1282
QY 1145 AGAAGTGCCTGCACCTCATCAAGCTGAAGATTTCTGTGTGAGCCTGGTGCATGTCAAAG 1204
Db 1283 AGAAGTGCCTGCACCTCATCAAGCTGAAGATTTCTGTGTGAGCCTGGTGCATGTCAAAG 1342
QY 1205 GCCGTGTGTGGTGGCTCTGGCGGACGGGACCCCTGGCCATCTTCCACCGTGTGAAGATG 1264
Db 1343 GCCGTGTGTGGTGGCTCTGGCGGACGGGACCCCTGGCCATCTTCCACCGTGTGAAGATG 1402
QY 1265 GCCAGTGGATCTGAGCAACTATCACTTAATGGACCTGGGCCACCGCACCACTCCATCC 1324
Db 1403 GCCAGTGGATCTGAGCAACTATCACTTAATGGACCTGGGCCACCGCACCACTCCATCC 1462
QY 1325 GCTGCATGGCTTTGTGTACGACCGCGTGTGTGTGGCTTACAGAAACAAGGTGCACGTCA 1384
Db 1463 GCTGCATGGCTTTGTGTACGACCGCGTGTGTGTGGCTTACAGAAACAAGGTGCACGTCA 1522
QY 1385 TCCAGCCCAAGACCATGACATAG-- 1408
Db 1523 TCCAGCCCAAGACCATGACATAGAGGCGAGTGGCGGCCCGGGAGGGAAGA 1582
QY 1409 ----- 1408
Db 1583 GGCTCCTGTGGCCAGCAGCTCTCCCGCATCTTCCATACGGAAGTCCAGGAGCCCTACCC 1642
QY 1409 -----AGAGTCAATTGACGCCACCGCGGGGAGAGCAGGTGGCGAGCTG 1458
Db 1643 ATGACTCCACAGAAGTCAATTGACGCCACCGCGGGGAGAGCAGGTGGCGAGCTG 1702
QY 1459 CGGTGGATCGCGATGGCGTATGGTGTGTCATCGCCTGGATCCACCTGAGGCTTAC 1518
Db 1703 CGGTGGATCGCGATGGCGTATGGTGTGTCATCGCCTGGATCCACCTGAGGCTTAC 1762
QY 1519 CATGCACACGACACCATCTACAGGACGTGGACATTGAGCCCTAGCTCAGCAAGATG 1578
Db 1763 CATGCACACGACACCATCTACAGGACGTGGACATTGAGCCCTAGCTCAGCAAGATG 1822
QY 1579 CTAGGCACTGGCAAGCTGGGTTTCTCTTCTGTAAGCATCAGGCGCTGTGTGCGGGC 1638
Db 1823 CTAGGCACTGGCAAGCTGGGTTTCTCTTCTGTAAGCATCAGGCGCTGTGTGCGGGC 1882
QY 1639 AGCGGGCTCTGGTGGGACCGGCAACGGAGTGTCTCTTCTGTAAGCATCAGGCGCTGTGTGCGGGC 1698
Db 1883 AGCGGGCTCTGGTGGGACCGGCAACGGAGTGTCTCTTCTGTAAGCATCAGGCGCTGTGTGCGGGC 1942
QY 1699 GTGTCTGTGACCGAGCCAGCTCTCTGGGCTCGGAGCCAAAGACATCCGCCACCTCT 1758
Db 1943 GTGTCTGTGACCGAGCCAGCTCTCTGGGCTCGGAGCCAAAGACATCCGCCACCTCT 2002
QY 1759 GGGAGGGCGCGCTCCCGGGGACATCATCAAGTGTATGGCGATGACAGTGCAGG 1818
Db 2003 GGGAGGGCGCGCTCCCGGGGACATCATCAAGTGTATGGCGATGACAGTGCAGG 2062
QY 1819 GCGGCCAGCAGCTTCATCCCTACTGTCTCATGCCCCAGGCCAGCTATGCTTCCATGGG 1878
Db 2063 GCGGCCAGCAGCTTCATCCCTACTGTCTCATGCCCCAGGCCAGCTATGCTTCCATGGG 2122
QY 1879 CACCGGATCCGCTGAAGTTCTTTGTCTCGGTGCGAGGAACGTGTGGCCACCTCAAT 1938
Db 2123 CACCGGATCCGCTGAAGTTCTTTGTCTCGGTGCGAGGAACGTGTGGCCACCTCAAT 2182
QY 1939 GCGAGTGTGTGACAGCCAGCGGCGGCTGGGCCAGCTGCCCTGCTCGGAGGTC 1998

Db 2183 GGCAATGTCTGTGACAGCCAGCCAGAGGCGCTTGGGCCAGCTGCCCTCGAGGTC 2242
QY 1999 GAGGGCCAGAAAGCTGCGGAACGCTGTGGTGTCTGAGCGGGGAGGGCTACATCGACTTC 2058
Db 2243 GAGGGCCAGAAAGCTGCGGAACGCTGTGGTGTCTGAGCGGGGAGGGCTACATCGACTTC 2302
QY 2059 CGCAATTTGGAGACGAGAGGACGACGAGACGAGAGGAGGGGCGCATGAGCCAGGTG 2118
Db 2303 CGCAATTTGGAGACGAGAGGACGACGAGACGAGAGGAGGGGCGCATGAGCCAGGTG 2362
QY 2119 AAGCCCGTCTGTCCAAAGCAGAGCGCAGTCAATCATCTGTGTGGCAGGTGCTTCAACC 2178
Db 2363 AAGCCCGTCTGTCCAAAGCAGAGCGCAGTCAATCATCTGTGTGGCAGGTGCTTCAACC 2422
QY 2179 CCCGAGTGA 2187
Db 2423 CCCGAGTGA 2431

RESULT 12
ABA09028
ID ABA09028 standard; cDNA; 4688 bp.
XX
AC ABA09028;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human secreted protein homologue-encoding cDNA, SEQ ID NO:804.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytosstatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003800.
XX
PR 03-FEB-2000; 2000US-00496914.
XX
PR 27-APR-2000; 2000US-00560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-457740/49.
XX
XX P-PSDB; ABB11784.
XX
PT Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
XX
XX
PS Claim 1; Page 710-712; 1963pp; English.
XX
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which

CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC providing an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention

XX Sequence 4688 BP; 1438 A; 931 C; 1163 G; 1156 T; 0 U; 0 Other;

Query Match 13.2%; Score 486.2; DB 4; Length 4688;
 Best Local Similarity 58.5%; Pred. No. 3e-85;
 Matches 1191; Conservative 0; Mismatches 753; Indels 93; Gaps 16;
 1973 GGTCAAGGTGAAATAAGATGAAATTTACCTGTGCTGTCTATCTCAGACCTCTGGAT 2032
 214 GCGCAGGAGGACCGGATGAAGAACGTGCGGTGCGCGGTGATCTGCGCCCTCTGGTG 273
 274 GAGAGGACCCACCATGAGCTGTGTGTCGCGGGCGTCAACTGAGCGGGTGGAGG 333
 2033 GAAAGATATACATCAATGAAGCTGTGTGCTGTGTTGGAGTCAATTTATCTGTTGGAG 2092
 334 CCCAATGAGGACGACCTCGGAATGGAG-----TCAAGCCAGCGCCAGCGCGC 381
 2093 ACCAGAGATGGTGTCTGTGTTGGAGCAAGTGATTTTACAAGATGTTGCTGTTG 2152
 382 GATCCCCCTGACCTGGACCGGAGAGACGGCGAGCCCAAGAGCGCCACAGCTCTCC 441
 2153 GATACAGAAAGGAGTAAACAGCGAAGTGCCTCTCAGAGTAGTATTAGATAAGTTAGATCAG 2212
 442 GAG---AAGAGAGCCCAAGAGCTCCCTGAAATGGAGCGCCACCTCCAGCGGGTGGG 498
 2213 GAATTTAAGAAACAGCAGAGGAGTTAAATAATCAAGAGAATTATCCAGTCTAGTTGG 2272
 499 ATCCCTGACGACCCCTGACCAACAGCAAGGTGGTGATCATCGACGCCCAACAGCGGGC 558
 2273 ATCTGTACAGCACTCATTCGGCTACAAAAGTTCTTTATTATTGATGCTGTTCAACTGGC 2332
 559 ACGGTGGTGGACAGTTACCGTCTGCAAGCGGCACGTGCTGTGATCTCCAGCATCCCC 618
 2333 AACATCCTAGACAGTTTCACTGTTTGAACCTCTCATGTTCTGTGCAATTGCAAGTGTGCCA 2392
 619 GCGGCCAGCAGACGACGATACCTCCCGG--GGAGATGTTCTCTGACAGCGACGTGAACCC 677
 2393 GGTGCACGAGAAACAGACTACCTTCGAGGAGAGATCTTTCAAAATCTGTCAGGTAGAC 2452
 678 AGAGGACC-----CGGCGCAGATGCGGTGCTG 705
 2453 AAAGCATCTTTATGTGGAAGTATGACAAGCAACAGCTCAGCAGACAGACAGACGCTGTGA 2512

QY 706 GCCGTATACCTGTGGGCTGTGC--CACCCGCTCAACAGTGCCTGGGAGCAACTGCT 763
 Db 2513 GGAGGCAATCAGTGGTGTGTTCTGACAGAGGTGTGACGGGAGTGCACCTTCCCT 2572
 QY 764 CTTCCCGAGGGGACAC--CCAGTGTAGACAAGGGGAGGGGAGGTGGCCACCATCGCC 822
 Db 2573 AGTACAAATGGTGTCTTCCAGTGATGATAAAACACACAGAAATGGAAGCAGAAATAGT 2632
 QY 823 AACGGGAAGTCAACCGTCCCACTCCACAGAGAGGCCACAGAGGCCACGAGGTGCCA 882
 Db 2633 GA--GTTGATGAAATGTTCCACACAGAGAAAGCACTGAAGCTACAGAGGG--A 2688
 QY 883 GACCCTGGGCGCAGCGAGCCAGACACACATTTGGCGCCCGGCGCTCTCACAGAGCAC 942
 Db 2689 ATGGGGGTGAGTGAAGACACAGTTGGACATCTCCAAATCTGGCGTCTACACAGAGCAT 2748
 QY 943 GTCTTCACTGACCAG-----CCCCGACCCCGTCTCTGGCCCCCAGCC 986
 Db 2749 GTCTTTACAGATCCTTTGGGAAGTTTCAATCCCAAGAGACCTCTCCCACTGTATCAGTC 2808
 QY 987 TGGCAGCGAAGACGGGCGCAGAGCCTGTGACAGCAGCAGCACACGGCA--GAGCCAGAGCC 1043
 Db 2809 GAGCAATGACTCAGATGATATAAAGATCAAAATATCAGTACTGCAATGAACAGACTT 2868
 QY 1044 CAGCGGGACCCCAAGGAGCAGGAGCAGTGTGACCCCACTGTGGCTGGGAGGCCA 1103
 Db 2869 GGTGAGAGAAAGCCCAAGAAATGAGTAGTCTTTTACCAACTATGTGGCTGGAGTCA 2928
 QY 1104 GAACGGCTGGCTCTATGTGCACTGGCTGTGGCCAACTGGAGAAAGTCCCTGCATCCAT 1163
 Db 2929 AATGGCTGTTGTATGTCCATTCATCTGAGCCAGTGGAGAAATGCTCCATTCAT 2988
 QY 1164 CAAGCTGAAGATCTGTGCTGAGCCTGTGATCAAGAGCCGTGTGCTGGCTCT 1223
 Db 2989 TAAACTTAAAGATTCGATCTCAGTATTGTACATGAAAGGAATCGTGTAGTAGCCCT 3048
 QY 1224 GCGCAGCGGACCCCTGGCCATCTTCCACCGTGTGAAGATGCGCCAGTGGGATCTGAGCAA 1283
 Db 3049 GGCTGACGGCACCTTGCATCTTTCACAGAGGAGTGGATGGGAGTGGGATTTGTCAA 3108
 QY 1284 CTATCACTAATGAGACTGGGCGCCACCCGACCTCCATCCGCTGCATGGCTGTGTGTA 1343
 Db 3109 CTATCACTCTTAGACCTTGGACGCGCTCATCATTCATTCGCTGATGACTGTGTGTA 3168
 QY 1344 CGACCGCTGTGTGCTGCTCAAGAAACAAAGTGCAAGTGCACGCTCATCCAGCCCAAGACCATGCA 1403
 Db 3169 TGACAAAGTCTGTGTGGCTATAGGAACAAATCTATGTGTGAGCCAAAGGCCATGAA 3228
 QY 1404 GATAGAGAGTCAATTTGACGCGCCACCCGCGGGAGAGCCAGGTGCGGACGTGGCGTG 1463
 Db 3229 AATAGAGAAATCTTTTGTGATGCAATCCAGAGAAAGAGAGCCAAAGTGCAGAGTTCG 3288
 QY 1464 GATCGGAGTGGCTATGGGTGCTCATCGCCCTGCACTCCACCTGAGGCTCTACCATGC 1523
 Db 3289 GGTGGGGATGCGGTGCGGTCTCCATTCGCTTGGATTCTACGCTCCGCTCTATCATGC 3348
 QY 1524 ACACAGCACCAGCATCTACAGGACGTGGACATTTAGCCCTACGTGAGCAAGATGTAGG 1583
 Db 3349 ACACATTTAATCAATCTACAGATGTGACATTTGAGCTTATGTAAGCAAAATGTTAGG 3408
 QY 1584 CACTGCAAGCTGGGTTTCTCTTGTGATGCAATCAAGCGCCCTGCTTGTGCGGGAGCGG 1643
 Db 3409 TACTGGAAGTGGGCTTCTCTTTTGTGAGAAATACAGCTCTTATGCTTCTTGTAAATCG 3468
 QY 1644 GCTCTGGGTGGGACCGGCAACGGAGTGTCTATCTCATCCCCCTGACAGAGACTGTGT 1703
 Db 3469 TTTGTGGGTGGGACAGGAAATGGTGTCTATCTCCATCCCAANTGACAGAAACCGTAT 3528
 QY 1704 CTTGACCCGAGGCGACCTCTGGGGCTCCGAGGCCAATAAGACATCCCCCACTCTGGGGA 1763
 Db 3529 CTTCCACAGGACGTTTACTGGGGCTGAGGGCAATAAACCTCAGGTGTACAGGAAA 3588

QY 1764 GGGGCGCCGTCCTCCGGGGGATCATCCACGTTGTATGGCATGACAGCAGTGCACA---GGGC 1820
 Db ||||| T-----CGTCTCGAAGTGTAAATCCGTGTATATGGTGTATGATGAAACAGTGTAAAGTGAC 3642
 QY 1821 GGGCAGCAGCTTCATCCCTACTGCTCCATGGCCCGAGGCCAGCTATGCTTCCATGGCA 1880
 Db ||||| TCCAGGACATTTATACCTTATTTGTTCAATGGCAGCATGCACAGCTTGTCTTCCATGGCA 3702
 QY 1881 CGCGATGCGGTGAAGTTCTTTCTCTCGGTGCCAGGGAACGCTGCTGGCCACCCTGAATGG 1940
 Db ||||| CCGGATGCTGTGAATTTCTTTGGCAGTCCCGAGTCAAGTCAATCAGCCCAAAAGTAG 3762
 QY 1941 CAGTGTGTCGACAGCCCGAGGCGCTTGGCCAGCTGCCCTGCCGAGGTGCA 2000
 Db ||||| CAGT-AGTGGCAGGATCTGACGGGTGCAAAAGCGAGGGGCCATCTGCACAGGAGCTGG 3821
 QY 2001 GGGCCAGAAAGC---TSCGGAACGTGCTGTGCTGACGAGGAGGGCGGAGGGGCTACATCGACTT 2057
 Db ||||| TAGTCAGAGCCCTTGAAGTCTATGCTTGTCTATCATCAGTGGAGAGAGGGCTACATCGACTT 3881
 QY 2058 CCGCATTTGGAGCGGAGAGGACGACGAGAGGAGGGCGGAGGGGAGGGGCTACATCGACTT 2117
 Db ||||| CCGAATGGGTGATGAAG-----GTGGAGATCAGAACTTCTTTGGAGAGGATCTTCCACTT 3936
 QY 2118 GAAGCCCTGCTGTCTCAGGCGAGGCGGAGTCAATCATCATCATCATCATCATCATCATCATCAT 2174
 Db ||||| GAACCTTCTG-TCAACAAAGCAAGAGGAGTCACTTGTATGTGTGCAAGTGTATGTA 3992

RESULT 13

ABX63549

ID ABX63549 standard; cDNA; 4286 BP.

AC ABX63549;

XX

XX

DT 26-FEB-2003 (first entry)

XX

XX

DE Human cDNA #549 differentially expressed in activated vascular tissue.

XX

KW Human; gene; ss; vascular tissue; cytostatic; hypertensive; antidiabetic;

KW hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;

KW gene therapy; vascular disease; cancer; coronary; artery disease;

KW hypertension; diabetes; pre-eclampsia; restenosis;

KW ischaemia-reperfusion injury; stroke.

XX

OS Homo sapiens.

XX

XX

FN US2002137081-A1.

XX

XX

PD 26-SEP-2002.

XX

XX

PF 08-JAN-2002; 2002US-00044090.

XX

XX

PR 28-JUL-2000; 2000US-0222469P.

XX

XX

PR 08-JAN-2001; 2001US-0260483P.

XX

XX

PA (BAND/) BANDMAN O.

XX

XX

PI Bandman O;

XX

XX

DR WPI; 2003-110597/10.

XX

XX

PT Combination for diagnosing, staging, treating, or monitoring the

PT progression of treatment of a vascular disease, e.g. atherosclerosis;

PT comprises several cDNAs that are differentially expressed in activated

PT vascular tissue.

XX

XX

PS Claim 1; Page; 18pp; English.

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Sequence 4286 BP; 1323 A; 781 C; 1025 G; 1157 T; 0 U; 0 Other;

Query Match

Best Local Similarity 12.9%; Score 477; DB 7; Length 4286;

Matches 1154; Conservative 56.5%; Pred. No. 1.8e-83;

Matches 1154; Conservative 0; Mismatches 805; Indels 84; Gaps 11;

QY 214 GGCAGGAGGACACCGGATGAAGAACGTCGCGGTGCGCGGTGCTACTGCGCGCTCTGCTG 273

Db ||||| GGTCAAGGTGAAATTAAGATGAAATTTACTGTGCTGTCTATCTCAGACCTTGAT 1030

QY 274 GAGAGGAGCCCGATGAGCTGTGTGTCGCGGCGGTCAACCTGAGCGGGTGGAGG 333

Db ||||| GAAAAGATACATCAATGAAGCTGTGTGCTGTGAGTCAATTTATCTGTGGAG 1090

QY 334 CCCAATGAGGACGACGCTGGGAATGGAG-----TCAGCCAGCGCCAGCGCC 1149

Db ||||| ACCAGAGATGTTGTTCTGTTTGGAGCAAGTGATTTTACAAGGATGTTGCTGTTG 1150

QY 382 GATCCCTGACCTCGACCGCGAAGGAGAGCGCGAGCGCCCAAGCGCCACACGCTCCCC 1209

Db ||||| GATACAGAGGACGAGTAAACAGCGAAGTGCTCTCAGAGTAGTTTAGATAAGTTAGAT 1210

QY 442 GAG---GAGAGAGGCGCAAGGAGCTCCCTGAATGAGCGCCACCTCCAGCGGGTGTGG 1269

Db ||||| GAACTTAAGGAACAGCAGAGAGGAGTTAAAAAATCAAGAGAATTTATCCAGTCTAGTTGG 1270

QY 499 ATCTGACACGACCCCTGACCCAGCAGAGTGGTGTATCATCGACGCCAACCAGCGGGC 1329

Db ||||| ATCTGTACAGCACTCATTCGGCTCAAAAGTTCTTATTATTGATGCTGTTCAACCTGGC 1330

QY 559 ACGTGTGGTGGACCAAGTTCACTGTCGACGCGCAGCTGTGTGTCATCTCCAGCATCCCC 1389

Db ||||| AACATCCTAGACAGTTCCTGTTTGCACCTCTCATGTTCTGTGTCATTCAGATGTCGA 1390

QY 619 GCGGCCAGGACAGCGCACTACCTCCCGG-GGAGATGTTCTCGGACAGCGAGTGAACCC 1449

Db ||||| GGTGACAGGAAACAGACTACCTCGAGGAGAGATCTTTCAAGATCTGTCAGGTAGAC 1450

QY 678 AGAGGACC-----CGGGCCAGATGCGGTGCTG 705

Db ||||| AAGCATCTTTATGTGGAAGTATGACAAGCAACAGCTCAGCAGAGACAGACCTGTGA 1510

QY 706 GCGGTATCACCTGTGGGTGTGTCACCCGCTGCAAGTGCCTGCGGAGCACTGCTCC 1569

Db ||||| GGAGGATCACAGTGTGTTGTTCTGACAGAGTGTGACGGAGCTGCCCTTCCCT 765

QY 766 TCCCGAGGGGACACCCAGTGTAGACAGGGGCGAGGGGTGGCCACCATCGCAAC 825

Db |||||

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

1630 AGTACAAATGGTCTCTCCAGTGTATGATTAACACCCAGCAATGGAAGCAGAAATAGT 1689 Db
826 GGGAAAGTCAACCCGTCCTCCAGTCCACAGAGAGGCCACAGAGGCCACGAGAGTCCAGAC 885 Qy
1690 GAGTTGATGAATGTTCCACAGCAGCAAGAGCACTGAAGCTACAGAAGGGA--AT 1746 Db
886 CTTGGGCCCCAGGAGCAGAGACAGCCACATTTGGGCCCCGGGCTCTCACAGAGCAGTC 945 Qy
1747 GCGGGGTGAGTGAAGACAGAGTGAATCTCCCAAACTGGCGTCTACACAGAGCATGTC 1806 Db
946 TTCACTGACCC-----AGCCCGACCCCGCTCTCTGGCCCCCGACCTGGC 990 Qy
1807 TTACAGATCTTTGGGAGTTCAATCCAGAGACCTCTCCAGTGTATCACTCGAGC 1866 Db
991 AGCAGAAACGGGCCCCAGAGCCTGACAGCAGCAGCAGCCGCA---GAGCCAGAGCCGAGC 1047 Qy
1867 AATGACTCAGATGATATAAAGATCAATATCAGTACTGCCAAATGAACAAGACTTGGTG 1926 Db
1048 GGGACACCCAGGAGCAGGAGCAGAGTCTCCACCCACATGCTGGCTGGGAGCCGAGAC 1107 Qy
1927 AGAGAAGAGCCAGAAATGAGTAGTCTTTTACCAACTATGTGGCTTGGAGCTCAAAAT 1986 Db
1108 GGCTGGCTCTATGTCACCTCGGCTGTGCCAACTGGAAAGAGTGCCTGCACTCCATCAAG 1167 Qy
1987 GGTGTTTGTATGTCATCTCTAGTATGATAGCAGTGGAGAAATGTCCTCATTTAA 2046 Db
1168 CTGAAGGATTTGCTGAGCCTGGTGCATGTCAAAGCCGCTGTCTGCTGGCTCTGGCG 1227 Qy
2047 CTTAAAGATCGATTCTCAGTATTGTACAGCTGAAGGAATCGTGTAGTAGCCCTGGCT 2106 Db
1228 GACGGACCTCGGCATCTTCCACCGTGTGAGTGGAGTGGGATCTGAGCAACTAT 1287 Qy
2107 GACGGACCTTGCATCTTTTACAGAGAGTGGGATGGGAGTGTGTCAAACTAT 2166 Db
1288 CACCTAATGGACCTGGCCACCCGACCACTCCATCCGCTGCATGGCTGTGTGTACGAC 1347 Qy
2167 CACCTCTTTAGACCTTGGACGGCTCATATTCATCCGTTGATGACGTGTGTATGAC 2226 Db
1348 CGCGTGTGTGTGTGTACAGAACAAAGTGCAGTATCCAGCCCAAGCACTGAGATA 1407 Qy
2227 AAGTCTGTGTGTGTATAGGAACAAATCTATGTGTGACCCAAAGGCCATGAATA 2286 Db
1408 GAGAGTCAATTTGACGCCACCCGCGGAGAGCCAGTGGCGAGCTGGCGTGGATC 1467 Qy
2287 GAGAAATCTTTTATGATCATCCAGAGAGGAGCCAGTGGCAGCTGTGGTGGTG 2346 Db
1468 GCGGATGGGTATGGGTCTCCATCCGCTGGATCCACCTCGAGGCTTACCATGACAC 1527 Qy
2347 GGGGATGGGTGTGGGTCTCCATTCGTTGGATCTACGCTCCGTCTCTATCATGACAC 2406 Db
1528 AGCACCAGCATCTACAGCAAGTGGACATGAGCCCTCAGTACAGCAAGATGCTAGCACT 1587 Qy
2407 ACTTATCAACATCTACAGGATGGACATGAGCCCTTATGTAAGCAAAATGTTAGTACT 2466 Db
1588 GCGAGGTGGTTCCTTCCTGTAAGCATCAGCCCTGCTGTGCGGGCAGCCCGCTC 1647 Qy
2467 GGAATACTGGGCTCTCTTTTGTGAGATTTACAGCTCTTATGGTGTCTGTATCTGTTG 2526 Db
1648 TGGGTGGGACCGGCAACGAGTGGTGCATCTCCATCCCTGTACAGAGACTGGTCTGCT 1707 Qy
2527 TGGGTGGGACAGGAATGTGTCTATCTCCATCCCATTTGACAGAAACCGTAATCCTC 2586 Db
1708 CACCGAGGCGAGTCTGGGGCTCCGAGCCAAATAGACATCCCCCCTCTGGGGAGGC 1767 Qy
2587 CACCAGGAGCTTTACTGGGGCTGAGGGCAATAAACCCTCAGGTGTACAGGAAT--- 2643 Db
1768 GCCCGTCCCGGGGCATCATCCAGTGTATGGCGATGACAGCATGTGACA---GGGCGGCC 1824 Qy
2644 ---CGTCTCGAAGTGAATCCGTGTATATGGTATGAACAGTGAATAGTACTCCA 2700 Db
1825 AGCAGCTTCATCCCTACTGCTCCATGCGCCAGCCAGCCAGCTATGCTTCCATGGCAGCCG 1884 Qy
2701 GGGACATTTATACCTATTGTTCAATGGCAGATGACAGCTTTGCTTCCATGGGACCCGG 2760 Db

1885 GATGCGTGAAGTCTTTTGTCTCGGTGCCAGGGAACGTGCTGCCACCCCTGAATGGCAGT 1944 Qy
2761 GATGCTGTGAATCTTTTGTGGCAGTCCAGGTCAAGTCAATCAGCCACAAAGTAGCAGT 2820 Db
1945 GTGCTGGACACCCAGCCGAGGCGCCCTGGGCCAGCTGCCCTCGGAGTCCGAGGGC 2004 Qy
2821 AGTGGCAGCGATCTGACGGGTGACAAAGCAGGGCCATCTGCACAGGAGCCCTGGTAGTCAG 2880 Db
2005 CAGAACTGCGGAACGTCGTGCTGAGCGCGGGAGAGGCTACATCGACTTCCGCATT 2064 Qy
2981 ACGCCCTTGAAGTCTATGCTTTGTCATCATGTTGGAGGAGGGCTACATCGACTTCCGAATG 2940 Db
2065 GGAGACGAGAGGACGACGAGACGAGGAGGCGGCCAGCGGAGCATGAGCCAGGTTGAAGCCC 2124 Qy
2941 GGTGATGAAG-----GTGGAGATCAGAACTTCTTGGAGAGGATCTTCCACTTGAACCTT 2995 Db
2125 GTGCTGTCCAAAGCAGAGCGCAGTCAATCATCTGTGTGGCAGGTGCTCTACACCCCGAG 2184 Qy
2996 CTG-TCACCAAGCAGAAAGGAGTCACTTGATGTGTGGCAAGTATGTATGGCAATGAG 3054 Db
2185 TGA 2187 Qy
3055 TGA 3057 Db

RESULT 14
ADA45191 ID ADA45191 standard; cDNA; 4667 BP.
XX
AC ADA45191;
XX
DT 20-NOV-2003 (first entry)
XX
DE Murine JLP coding sequence.
XX
KW Cytostatic; Gene therapy; murine; scaffolding protein; JLP;
KW JNK-associated Leucine zipper protein; MEK kinase 3; MEK3;
KW MAP kinase kinase 4; MKK4; c-Jun NH2-terminal kinase; JNK;
KW p38 MAP kinase; MAPK; c-Myc; MAX; apoptosis; cancer; gene; ss.
XX
OS Mus musculus.
XX
FH Key
FT CDS 163..4086
FT /tag= a
FT /product= "Murine JLP"
XX
PN WO200306652-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003355.
XX
PR 05-FEB-2002; 2002US-0354377P.
XX
PA (UTEM) UNIV TEMPLE.
XX
PI Lee CM, Dhanasekaran N, Reddy PE;
XX
XX WPI; 2003-731487/69.
XX
DR P-PSDB; ADA45192.
XX
XX New scaffolding nucleic acid sequences, designated as JLP, useful for
XX modulating apoptotic response in a cell, and thus for treating metastatic
XX cancer.
XX
XX Claim 10; Page 79-80; 102pp; English.
XX
XX The present invention relates to novel human and murine scaffolding
XX proteins, JLP (for JNK-associated Leucine zipper protein, ADA45190 and
XX ADA45192). JLP tethers MEK kinase 3 (MEK3), Mitogen-Activated Protein
XX (MAP) kinase kinase 4 (MKK4), c-Jun NH2-terminal kinase (JNK), p38 MAP

CC kinase (MAPK), c-Myc and MAX into a signalling module which controls the
CC apoptotic response. JLP therefore functions as a signalling conduit to
CC transmit extracellular signals to the nucleus through MEK3-MKK4-
CC JNK/p38/MAPK/c-Myc/MAX signalling module. The JLP sequences are useful
CC for modulating apoptotic response in a cell, and thus for treating
CC metastatic cancer. The present sequence is the coding sequence for murine
CC JLP.

XX
SQ Sequence 4667 BP; 1376 A; 1024 C; 1201 G; 1066 T; 0 U; 0 Other;

Query Match 12.0%; Score 442.2; DB 8; Length 4667;

Best Local Similarity 56.5%; Pred. No. 1.1e-76;

Matches 1159; Conservative 0; Mismatches 763; Indels 129; Gaps 13;

214 GGCAGGAGGACACGCGGATGAAGACGTCGCGGTGCGGTGCTACTGCGCGCTCTGTG 273

DB 2098 GGTCAAGGGGAACTAAGATGAATAATTTACCTGTGCTCTATCTCAGACCCTAGAC 2157

QY 274 GAGAGGACCCACCATGAAGCTGTGTGTCGCGCGGCGTCAACCTGAGCGGTTGAGG 333

DB 2158 GAAAGAGATGCGTCTATGAAGCTGTGTGTCGTGTAGGAGTCAATTTATCTGTGSCAAG 2217

QY 334 CCCAATGAGGACGACGCTCGGAATGGAG-----TCAAGCCAGCGCAGGCCGC 381

DB 2218 ACAAGAGATGGCGGTCTGTGTGAGCGAGTGATTTTACAAGGATTTGCTGTTTG 2277

QY 382 GATCCCTGACCTGCGACCGCGAAGAGACGCGGAGCCCAAGAGCGCCACACGCTCCC 441

DB 2278 GACACTGAAGGCAGTAAACAGCGGAGTGGTCTCAGAGTAGTTTACAGAGTTGATCAA 2337

QY 442 GA--GAAGAAGAGCCAGGAGCTCCCTGAAATGGAAGCCACCTCCAGCGGGTGG 498

DB 2338 GAACTCAAGGACACAGCAGAAAGATTTAAAAATCAAGAAGAGTTTCCAGTCAGTCTGG 2397

QY 499 ATCTGTGACGACCTGACACACGAGAGTGGTGTATCATCGAGCCCAACAGCGCGGC 558

DB 2398 ATCTGTACAGACCCACTCACTCAAGAGTCACTCATCATTTGCTGTTTCAGCCTGCG 2457

QY 559 ACGGTGFGGACGATTCACCGTCTGCAACGCGACGCTGTGTGTCATCTCCAGCATCCC 618

DB 2458 AACATCCTAGATAGTTTCACTGTTTGCAATTTCTCATGTTCTGTGATTCAGAGTCCCA 2517

QY 619 GCGGCCAGCAGGACCTACCTCCCGG--GGAGATGTTCTGACAGCGACGTAACCC 677

DB 2518 GGAGCTCGGGAAACAGACTACCTCCTCGAGGAGAAACTTTCTGAATCTGGTCAGTAGAC 2577

QY 678 AAGGACCCGCGCGCAGATG-----GGTGTGCT 705

DB 2578 AAAGCGTATTGTGGAAGCATGACCATTAACAGCTCCGAGAGATGGACAGCTTGTG 2637

QY 706 GCCGGTATACCTGGTGGGCTGTGCCACCGCTGCAACGTTGCCGCGAGCAACTGCTCC 765

DB 2638 GGAGGCATCAAGTGTGGTGTCTCCACAGAAGACTGACAGAGCTGCCACTTCCCA 2697

QY 766 TCCCGAGGGACACCCAGTGTAGACAAGGGGAGGGGAGGTGGCCACCATCGCCAAC 825

DB 2698 AGTACCAACGGTGTCTCTCGTGAITTTGAAAGCCACAGAAATGGAATCTGAAATAGC 2757

QY 826 GGGAGAGTCAACCGCTCCAGTCCACAGAGAGGSCCAGAGGCCACGAGGTGCCAGAC 885

DB 2758 GAAGTTGATGAATATTCCAAAGCAGAGAAGAACTGAAGCCACAGAGG---GCAAT 2814

QY 886 CTGGGGCCAGCGAGCCAGAGACAGCCACATTTGGGCGCGGCGCTCTCAGAGACAGTCC 945

DB 2815 GCAGGGTCCACTGAAGACATGTGGACATCTCCAGCTGGCGTGTACAGAGCATGTG 2874

QY 946 TTCACTGACCCA-----GCCCGGACCCGTCCTCTGTGGCCCCAGCCTGGCAGCGAGAA 998

DB 2875 TTTCAGATCACTGGGAGTTCAGATCCCAAGAACCTCTCCCGAGTGTTCAGTCAAGT 2934

QY 999 CGGGCCAGAGCCCTGACAGCAGCAGCA-----CGGCCAGAGCCAGGCCAGC 1047

DB 2935 AATGACTCAGATGTACAAAGATCAGATATCAGTATTGTCAAATGAACAGACCTTGC 2994

QY 1048 GGGGACCCACCGGAGCAGGACGACGAGTGTGTCACCCACCATGTGGCTGGAGCCCAAC 1107

DB 2995 AGAGAAGAAGCTCAGAAAATGAGTAGTCTTTTACCAACCATGTGGCTTGGAGTCAAGAT 3054

QY 1108 GCGTGGCTCTATGTGCACTCGGCTGGCCCACTGGAAGAAGTGCCTGCACTCCATCAAG 1167

DB 3055 GCGTGTGTATGTCCACTCATCTGTAGCCCACTGGAGGAAATGTCTCCATTCATTAAAG 3114

QY 1168 CTGAAGGATTTCTGTGCTGAGCCCTGTGTCATGTCCAAAGCCGCTGTGGTGGCTCTGGCG 1227

DB 3115 CTCAAAGACTCAATACTCAGTATGTACATGTGAAGAAGAAATGTACTAGTGGCCCTGGCG 3174

QY 1228 GACGGACCCCTGGCCACTTTCCACCGTGTGAAGATGCGCAGTGGGATCTGAGCAACTAT 1287

DB 3175 GATGGCACCCCTTGCATCTTCCACAGAGAGTTGATGGACAATGGATCTGTCAAACTAT 3234

QY 1288 CACCTAATGGACCTGGGCCACCGCACCACTCATCTCCGCTGCATGCTGCTGTGTACGAC 1347

DB 3235 CACCTTTTAGACCTTGGAGCTCTCATCTTCCATACGATGCATGCTGTGTCATGAC 3294

QY 1348 CCGCTGTGTGTGCTTACAAGAACCAAGGTGCAAGTCTACCCAGCCCAAGACCATGAGATA 1407

DB 3295 AAAGTCTGTGTGGCTATAGSAAACAATCTATGTGTTCAACCAAGGCTATGAGATA 3354

QY 1408 GAGAAAGTCAATTGACGCCCAACCGCGCGGAGAGCCAGGTGCGGAGCTGGCTGGATC 1467

DB 3355 GAGAAATCATTTGATGTCACACCCAGGAAGAGAGCCAAAGTACGAGCTTGCATGGGTG 3414

QY 1468 GSCGATGCGGTATGGTGTCCATCCGCTGGACTCCACCTGAGGCTCTTACCATGCACAC 1527

DB 3415 GTGATGTGTATGGGTCTCCATTCGTTTGGATTCCAGCTCCGCTCTATCATGCACAT 3474

QY 1528 ACGCACCAAGCTCTACAGAACGTGACATTTAGCCCTAGCTCAGCAAGATGCTAGGCACT 1587

DB 3475 ACATACCAACATCTCCAGAGTGTGATATTGAGCCTTATGAAAGCAAAATGTTAGTTACT 3534

QY 1588 GCGAAGCTGGGTCTCTCCTCGTAGCATACGCGCTCTGTTGTCGCGGAGAGCCGGTCTC 1647

DB 3535 GGAAGACTGGGTCTCTCTTCGTGAGGATAACAGCCCTTATGCTCTTGCATGTTG 3594

QY 1648 TGGGTGGCACCGCAACGGAGTGGTCTCATCTCCATCCCTTGACAGAGACTGTGGTCTGT 1707

DB 3595 TGGGTGGGAGCTGGAATGTTGTCATTATCTCCATCCCACTGACAGAAACA----- 3645

QY 1708 CACGAGGCCAGTCTCTGGGCTCCGAGCCAAATGAACATCCCCCACTCTGGGAGGGC 1767

DB 3646 -----AATAAACCTCAGGACTCCAGGTAA----- 3671

QY 1768 GCCCGTCCCGGGGACATCATCCAGTGTATGGCGATGACAGCAGTGAAGGCGGCCA-- 1825

DB 3672 --CCGTCCTGGAAGTGAATCCCGTGTATATGGTGATGAGAACAAGTGAAGTACTCCA 3729

QY 1826 -GACGCTTCATCCCTACTCTCATGGCCCGAGCCAGCTATGCTTCCATGGGACCCG 1884

DB 3730 GGCACATTTAACCCTACTGTTCAATGGGCACATGSCACAGCTTTGTTTCCAGGGCACCCG 3789

QY 1885 GATGCCGTGAAGTCTTTGTCGCTGCCAGGCAACGTCGTGGCCACCTCTGAATGGCAGT 1944

DB 3790 GATGCTGTGAATTTCTTTGTGGCAGTCCCAAGGTCAAGTCAATTAGCCCAAAAGTAGCAGT 3849

QY 1945 GTCTGACAGCCAGCGAGGCCCTGGGCCAGCTGCCCTCTCGCTCGGAGTTCGAGGGC 2004

DB 3850 GGTGGGCGAGATCTTACAGCTGAC---AAGGCAGGGTCTATCTGCACAGAGCCAGTAGC 3906

QY 2005 CAGAACG---TGCGGAACGTGTGTGTGTCAGCGGCGGGAGGGCTTACATCGACTTCCGC 2061

DB 3907 CAGACCCCTTGAAGTCCCATGCTTGTTCATCAGTGGAGAGAGGGCTACATTGACTTCCG 3966

QY 2062 ATTGAGAGCGGAGAGGACGACGAGAGCGAGGGGCGCAGGGGACATGAGCCAGGTGAAG 2121

DB 3967 ATGGGTGATGAAGGTGGAGA-----ATCTGAATCTTCTGGGAGAGGATCTTCCACTGGAA 4020

QY 2122 CCGTGTCTGTCACAGGCGAGCGAGTCACATCATCTGTGTGGCAGGTGCTCTACACCC 2181
DB 4021 CCTTCAGTCCACCAAGCGGAGAGAGTCACTGATAGTGTGGCAAGTGTGGCAAT 4080
QY 2182 GAGTGAAGCTG 2192
DB 4081 GAGTGAAGCTG 4091
RESULT 15
ADA45189
ID ADA45189 standard; cDNA; 4675 BP.
XX AC
XX ADA45189;
XX DT 20-NOV-2003 (first entry)
XX DE Human JLP coding sequence.
XX KW Cytostatic; Gene therapy; Human ; scaffolding protein; JLP;
XX KW JNK-associated Leucine zipper Protein; MEK kinase 3; MEK3;
XX KW MAP kinase kinase 4; MKK4; c-Jun NH2-terminal kinase; JNK;
XX KW p38 MAP kinase; MAPK; c-Myc ; MAX ; apoptosis; cancer; gene; ss.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX CDS 79..4014
XX FT /*tag= a
XX FT /product= "Human JLP"
XX PN WO2003066652-A2.
XX PD 14-AUG-2003.
XX PF 04-FEB-2003; 2003WO-US003355.
XX PR 05-FEB-2002; 2002US-0354377P.
XX PA (UTEM) UNIV TEMPLE.
XX PI Lee CM, Dhanasekaran N, Reddy PE;
XX WPI; 2003-731487/69.
XX DR P-PSDB; ADA45190.
XX PT New scaffolding nucleic acid sequences, designated as JLP, useful for
XX PT modulating apoptotic response in a cell, and thus for treating metastatic
XX PT cancer.
XX PS Claim 1; Page 74-75; 102pp; English.
XX PS The present invention relates to novel human and murine scaffolding
XX CC proteins, JLP (for JNK-associated Leucine zipper Protein, ADA45190 and
XX CC ADA45192). JLP teaches MEK kinase 3 (MEK3), Mitogen-Activated Protein
XX CC kinase 4 (MKK4), c-Jun NH2-terminal kinase (JNK), p38 MAP
XX CC kinase (MAPK), c-Myc and MAX into a signalling module which controls the
XX CC apoptotic response. JLP therefore functions as a signalling conduit to
XX CC transmit extracellular signals to the nucleus through MEK3-MKK4-
XX CC JNK/p38/MAPK/c-Myc/MAX signalling module. The JLP sequences are useful
XX CC for modulating apoptotic response in a cell, and thus for treating
XX CC metastatic cancer. The present sequence is the coding sequence for human
XX CC JLP.
XX SQ Sequence 4675 BP; 1444 A; 926 C; 1162 G; 1143 T; 0 U; 0 Other;
Query Match 11.2%; Score 413.4; DB 8; Length 4675;
Best Local Similarity 55.0%; Pred. No. 4.8e-71;
Matches 1124; Conservative 0; Mismatches 796; Indels 123; Gaps 11;
QY 214 GCCCAGGAGGACGCGGATGAAGAACGTGCCGGTGGCGGTGTACTCCCGCTCTGGT 273
DB 2026 GGTCAGGTGAATAAGATGAATAATTTACCTGTGCTGTATCTCAGACCTCTGGAT 2085

QY 274 GAGAGGAGCCACCACATGAGCTGTGTGTGCGCGGGCTCAACCTGAGCGGTGAAG 333
DB 2086 GAAAAAGATACATCAATGAAGCTGTGTGTGCTGTGGAGTCAATTTATCTGTGGGAAG 2145
QY 334 CCCAATGAGGACGACGCTGGGAATGGAG-----TCAAGCCAGCGCCAGCGCG 381
DB 2146 ACCAGAGATGTTGTTCTGTTGTTGGAGCAAGTATATTTACAGAGATGTTGCTGGTTG 2205
QY 382 GATCCCTGACCTGCGACCGGAGAGAGAGCGGAGAGCCAGAGCGGCCACACGCTCTCCC 441
DB 2206 GATACAGAAGCGAGTAACACAGCGAAGTGCCTCTCAGAGTAGTTTAGATAAGTTAGATCAG 2265
QY 442 GAG---AAGAAGAGGCGCAAGAGCTCCCTGAATGAGCGCCACCTCAGCGCGGTGTGG 498
DB 2266 GAACCTAAGGAACAGCAGAGAGGAGTTAAAAAATCAAGAAGAAATATCCAGTCTAGTTGG 2325
QY 499 ATCTGACCGACACCTTGACCAACAGCAAGTGTGTATCATCGACGCCCAACAGCGCGGC 558
DB 2326 ATCTGTACCAGCACTCATTCGGCTACAAAGTTCTTATTTAGTGTCTTCAACCTGGC 2385
QY 559 ACGTGTGGACCAAGTTACCGCTGCAACGCGACAGTGTGTGTCATCTCCAGCATCCCC 618
DB 2386 AACATCCTAGACAGTTTCACCTGTTTGCACCTCTCATGTTCTGTCATGCAAGTGTGCA 2445
QY 619 GCGGCGACGACAGCGACTACCTCCCGG-GGAGATGTTCTTGGACAGCGACGTGAACC 677
DB 2446 GGTGACAGAGAAACAGACTACCTCGAGGAGAAGATCTTTCAGAATCTGGTCAGGTAGAC 2505
QY 678 AGAGGACC-----CGGGCGAGATGGGTGTG 705
DB 2506 AAAGCATCTTTATGTGGAAGTATGACAAAGCAACAGCTCAGCAGAGACAGACGCTGTTA 2565
QY 706 GCCGTATCACCTGCTGGGTGTGCCACCGCTGCAACGTCGCCGCGAGCAACTGTCTCC 765
DB 2566 GGAGGCATCACAGTGGTTGGTTTCTCGAAGAGTGTGACGGAGCTGCCACTTCCCT 2625
QY 766 TCCCGAGGGACACCCAGTGTAGACAAGGGGCGAGGGGTGGCCACCAATCGCCAACT 825
DB 2626 AGTACAATGTTGCTTCTCCAGTGATGATAAACCAACCAAGAAATGGAAGCAGAAATAGT 2685
QY 826 GGGAAAGGTCAACCGTCCAGTCCACAGAGAGGCCACAGAGCGCCAGCGAGTGCAGAC 885
DB 2686 GAGGTTGATGAATAATGTTCCAAACAGCAGAGAAGCAACTGAAGCTACAGAAGG---AAT 2742
QY 886 CTTGGGCGCAGCGAGCGACAGACAGCCACATTGCGGCCCGCGCTCTCTCAGAGACGCTC 945
DB 2743 GCGGGTCAAGTGAAGACACAGTGGACATCTCCCAACTGGCGTCTACACAGAGCATGTC 2802
QY 946 TTCACTGACCC-----AGCCCCGACCCCGTCTCTTGGCCCCCAGCCTGGC 990
DB 2803 TTTCAGATCTTTGGGAGTTCAGATCCAGAGAAGACTCTCCCGAGTGTATCAGTCGAGC 2862
QY 991 AGCGAAGACGGGCGAGAGCTGACAGCAGCAGCAGCAGCGCCA---GAGCCAGAGCCAGC 1047
DB 2863 AATGACTCAGATGCATATAAAGATCAAAATATCAGTACTGCCAAATGAAACAGACTTGGTG 2922
QY 1048 GGGGACCCCGGAGCAGCGACGACGCTGTGCAACCCACCATGTGGTGGAGCGCCAGAAC 1107
DB 2923 AGAAGAAGACCCAGAAATAGTAGTCTTTTACCAACTATGTGGCTTGGAGCTCAAAAT 2982
QY 1108 GGCTGGCTCTATGTGCACTGGCTGTGGCCAACTGGAAGAAGTGCCTGCACTCATCAAG 1167
DB 2983 GGCTGTTTGTATGTCTCCATTCATCTGTAGCCAGTGGAGGAAATGTCTCCATTCCATTA 3042
QY 1168 CTGAAGATTTCTGTGCTGAGCTGTGTCATGTCAAAGCCGCTGTGTGTGGTGGCTTGGCG 1227
DB 3043 CTTAAAGATTCGATTTCTCAGTATTTGACAGTGAAGGAAATCGTGTGTAGAGCCCTGCT 3102
QY 1228 GAGGGACCTCGCCATCTTCCACCGTGTGAAGATGGCCAGTGGGATCTGAGCAACTAT 1287
DB 3103 GACGGCACCTTGCATCTTTTCACAGAGAGTGTGATGGCAGTGGGATTTGTCAAACTAT 3162


```
QY 1288 CACCTAATGGACCTGGGCGCACCGCCACCACTCCATCGCTGCATGGCTGTGTGTACGAC 1347
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3163 CACCTCTTAGACCTTGGACGGCCTCATCAATCCATCGTTGCATGACTGTGGTACATGAC 3222
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1348 CGGCTGTGGTGTGGCTACAGAACAGAGTGCAGTCAATCCAGCCCAAGACCATGCAGATA 1407
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3223 AAGTCTGGTGTGGCTATAGAAACAAATCTATGTGTGAGCCAAAGGCCATGAAATA 3282
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1408 GAGAAGTCATTTGACGCCACCGCGGGGAGAGCCAGCTGCGGAGCTGGCGTGGATC 1467
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3283 GAGAAATCTTTGATGACATCCAGGAGGAGGCAAGTGGCAGCTTGGCTGGGTG 3342
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1468 GCGGATGGCGTATGGGTGTCATCCGCTGGACTCCACCTGAGGCTTACCATGCACAC 1527
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3343 GGGGATGGCGTGTGGGTCTCCATTCGCTTGGATTACGCTCGTCTCTATCATGCACAC 3402
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1528 AGCACCAGCATCTACAGAGCTGGACATTTAGCCCTAGCTAGTCHAGGAAGATGTAGGCAC 1587
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3403 ACTTATCAACATCTACAGGATGGGACATTTAGCCCTTATTAAGCAAAATGTTAGGTACT 3462
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1588 GGCAGCTGGGTCTCTCTTCGTACCATCGGCCCTCTTGTGCGGGCAGCCGGCTC 1647
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3463 GGAATACTGGGCTTCTCTTTGTGAAATTAAGCTCTTATGGTGTCTGTAAATCGTTG 3522
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1648 TGGGTGGGACCGGCAACGAGATGGTCTCTCCATCCCTGACAGAGACTGTGTCCTG 1707
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3523 TGGGTGGGACAGGAAATGTTGTCATTTATCTCCATCCCATTTGACAGAAAC----- 3572
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1708 CACCGAGGCCAGCTCCTGGGGCTCCGAGCCATAGACATCCCTCTGGGGAGGGC 1767
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3573 -----AAATAAACCTCAGGTGTACCAGGA 3597
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1768 GCCCTGCCGGGGCATCATCCAGCTGTATGGCATGACAGCAGTGACAGGGCGGCC--- 1824
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3598 AATCGTCTCGAAGTGTAAATCGTGTATATGTATGTATGATGAACAGTGAAGTACTCCA 3657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1825 AGCAGTTTATCCCTACTGCTCCATGGCCAGCCAGCTATGCTTCCATGGGCACCGC 1884
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3658 GGGACATTTATACCCATATTGTTCAATGGCAGATGCACAGCTTTGCTTCCATGGGCACCGG 3717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1885 GATCCGTGAAGTTCTTTGTCTCGTGTGCGGAGGACGTGTGGCCACCTGAAATGGCAGT 1944
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3718 GATGCTGGAATTTCTTTGTGGAGTCCCAAGTCAAGTCAATCAGCCCAAAAGTAGCAGT 3777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1945 GTGCTGGACAGCCAGCGAGGGCCCTGGGCCAGCTGCCCTCGAGTCCGAGGCTCAGGGC 2004
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3778 AGTGCGACGGATCTGACGGGTGACAAAGCAGGGCCATCTGCACAGGAGCCTGGTAGTCAG 3837
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2005 CAGAAGCTGCGGAACGTGCTGGTGTGAGCGCGGGAGGGCTACATCGACTTCGCAATT 2064
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3838 ACGCCCTTGAAGTCTATGCTTGTCTCATCAGTGGAGGAGGGCTACATCGACTTCCGATG 3897
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2065 GGAGACGAGAGGAGCAGCAGCAGGAGGAGGGCGGAGGAGCATGAGCGAGGTGAAGCCC 2124
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3898 GGTGATGAAG-----GTGGGAATCAGAACCTTTGGAGAGGATCTTCCACTTGAACCTT 3952
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2125 GTGCTGTCCAAGGAGGAGCGCAGTCAATCATCGTGTGGCAGGTGTCTACACCCCGGAG 2184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3953 CTG-TCACCAAAGCAGAAAGAGTCACTTGTATGTGGCAAGTGTATGGCAATGAG 4011
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2185 TGA 2187
Db |||||
QY 4012 TGA 4014
Db |||||
```

Search completed: August 24, 2004, 02:42:01
Job time : 1288 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	3484.6	93.2	4701	9	US-09-964-864-326	Sequence 3
2	486.2	14.3	4688	13	US-10-276-774-804	Sequence 8
3	477	12.9	4286	14	US-10-044-090-549	Sequence 9
4	360.6	9.8	3791	15	US-10-240-965-3	Sequence 3
5	242.4	6.6	1413	9	US-09-925-297-167	Sequence 1
6	241.8	6.5	310	13	US-10-085-783A-1790	Sequence 1
7	241.8	6.5	310	16	US-10-243-535A-1790	Sequence 1
8	226	6.1	554	16	US-10-264-237-807	Sequence 8
c	219.4	5.9	675	9	US-09-917-800A-965	Sequence 9
10	189	5.1	189	9	US-09-796-692-6852	Sequence 6
11	189	5.1	189	15	US-10-040-862-6852	Sequence 6
12	189	5.1	189	16	US-10-057-475B-6852	Sequence 6
13	189	5.1	189	16	US-10-154-884B-6852	Sequence 6
14	116.6	3.2	6289	9	US-09-954-455-819	Sequence 8

QY 240 CGTGCCTGCGGTGTAATGCTGCGCCCTCTGTGGTGGAGAGAACCCACCATGAAGCTGTG 299
Db 1263 CGTGCCTGCGGTGTAATGCTGCGCCCTCTGTGGTGGAGAGAACCCACCATGAAGCTGTG 1322
QY 300 GTGTGCGCGGGCGTGAACCTGAGCGGTGAGGCCCAATGAGACAGCGCTGGGAATGG 359
Db 1323 GTGTGCGCGGGCGTGAACCTGAGCGGTGAGGCCCAATGAGACAGCGCTGGGAATGG 1382
QY 360 AGTCAAGCCAGCGCCAGCGCGCGATCCCTGACCTGGGACCGCGAAGAGACGCGAGCC 419
Db 1383 AGTCAAGCCAGCGCCAGCGCGCGATCCCTGACCTGGGACCGCGAAGAGACGCGAGCC 1442
QY 420 AAGAGCGCCACACGCTCTCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
Db 1443 CAAGAGCGCCACACGCTCTCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1502
QY 480 CACTTCCAGCGCGGTGATCTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 539
Db 1503 CACTTCCAGCGCGGTGATCTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1562
QY 540 CGAGCCAAACAGCGCGGACAGGTGGTGGACAGTTCACCGTCTGCAACCGCGACGTGCT 599
Db 1563 CGAGCCAAACAGCGCGGACAGGTGGTGGACAGTTCACCGTCTGCAACCGCGACGTGCT 1622
QY 600 GTGCATCTCCAGCATCCCGCGGCGCAGCGACAGCGACTACCTCCCGGGAGATGTTCC 659
Db 1623 GTGCATCTCCAGCATCCCGCGGCGCAGCGACAGCGACTACCTCCCGGGAGATGTTCC 1682
QY 660 GGACAGCGACGTGAACCCAGAGGACCCGGGCGCAGATGGCGTGTGGCGGTATCACCCCT 719
Db 1683 GGACAGCGACGTGAACCCAGAGGACCCGGGCGCAGATGGCGTGTGGCGGTATCACCCCT 1742
QY 720 GGTGGCTGTGCAACCCCGCTGCAACGTGCGGAGCAACTGCTCTCCCGAGGGGACAC 779
Db 1743 GGTGGCTGTGCAACCCCGCTGCAACGTGCGGAGCAACTGCTCTCCCGAGGGGACAC 1802
QY 780 CCCAGTGTAGAACAGGGGCGAGGGGAGGTGGCCACCATCGCCACCGGAGGTCAACCC 839
Db 1803 CCCAGTGTAGAACAGGGGCGAGGGGAGGTGGCCACCATCGCCACCGGAGGTCAACCC 1862
QY 840 GTCCAGGTTCACAGAGGAGGCGCACAGAGGCGACCGAGGTGCGACACCTTGGGCGCAGCGA 899
Db 1863 GTCCAGGTTCACAGAGGAGGCGCACAGAGGCGACCGAGGTGCGACACCTTGGGCGCAGCGA 1922
QY 900 GCCAGAGACGCCACATTCGGCGCGGGCTCTCAAGAGACGCTTCACTGACCCAGC 959
Db 1923 GCCAGAGACGCCACATTCGGCGCGGGCTCTCAAGAGACGCTTCACTGACCCAGC 1982
QY 960 CCCGACCCCGTCTCTGGCGCCCGAGCTGGCAGCGAGAACGGGCGCAGAGCCTGACAGCAG 1019
Db 1983 CCCGACCCCGTCTCTGGCGCCCGAGCTGGCAGCGAGAACGGGCGCAGAGCCTGACAGCAG 2042
QY 1020 CAGCACCGCGCAGAGCGACAGCGCGGACCCCGAGGAGCAGGACGAGTGTCTG 1079
Db 2043 CAGCACCGCGCAGAGCGACAGCGCGGACCCCGAGGAGCAGGACGAGTGTCTG 2102
QY 1080 ACCCACCAATGTGGCTGGGAGCCAGAAACCGCTGGCTCTATGTGCACTGGCTGTGCGCAA 1139
Db 2103 ACCCACCAATGTGGCTGGGAGCCAGAAACCGCTGGCTCTATGTGCACTGGCTGTGCGCAA 2162
QY 1140 CTGGAGAGTGGCTGCACTCCATCAAGCTGAAGGATCTGTGCTGAGCCTGGTGATGT 1199
Db 2163 CTGGAGAGTGGCTGCACTCCATCAAGCTGAAGGATCTGTGCTGAGCCTGGTGATGT 2222
QY 1200 CAAAGCCCGTGTGCTGGTGGCTCTGGCGGACGGGACCTTGGCCATCTTCCACCGTGGTGA 1259
Db 2223 CAAAGCCCGTGTGCTGGTGGCTCTGGCGGACGGGACCTTGGCCATCTTCCACCGTGGTGA 2282
QY 1260 AGATGGCCAGTGGGATCTGAGGCACTATCACCTAAATGGACCTGGGCAACCCGACCACTC 1319
Db 2283 AGATGGCCAGTGGGATCTGAGGCACTATCACCTAAATGGACCTGGGCAACCCGACCACTC 2342
QY 1320 CATCCGCTGATGGCTGTGTGTAACGACCGCGTGTGGTGGCTGTGTAACGAAACAGGTGCA 1379

Db 2343 CATCCGCTGATGGCTGTGTGTAACGACCGCGTGTGGTGGCTGTGTAACGAAACAGGTGCA 2402
QY 1380 CGTCATCCAGCCCAAGACCATGAGATAGAGAGTCAATTTGACGCCACCCGCGCGCGGA 1439
Db 2403 CGTCATCCAGCCCAAGACCATGAGATAGAGAGTCAATTTGACGCCACCCGCGCGCGGA 2462
QY 1440 GAGCAGGTGCGGAGCTGGCTGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 1499
Db 2463 GAGCAGGTGCGGAGCTGGCTGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 2522
QY 1500 CTCCACCCCTGAGGCTCTACCATGACACACGACCCAGCATCTACAGACGTGAGCATTTGA 1559
Db 2523 CTCCACCCCTGAGGCTCTACCATGACACACGACCCAGCATCTACAGACGTGAGCATTTGA 2582
QY 1560 GCCCTACGTGAGCAAGATGCTAGGACCTGGGCAAGCTGGGTTCTCTCTGCTACGCATCAC 1619
Db 2583 GCCCTACGTGAGCAAGATGCTAGGACCTGGGCAAGCTGGGTTCTCTCTGCTACGCATCAC 2642
QY 1620 GGCCCTGCTTCTGCGGGGACGCGGCTCTGGGTGGGACCCGCGCAACGAGGTGCTCATCTC 1679
Db 2643 GGCCCTGCTTCTGCGGGGACGCGGCTCTGGGTGGGACCCGCGCAACGAGGTGCTCATCTC 2702
QY 1680 CATCCCTCTGACAGACTGTGTCTGCAACGAGGCGAGCTCTCTGGGGCTCCGAGCCAA 1739
Db 2703 CATCCCTCTGACAGACTGTGTCTGCAACGAGGCGAGCTCTCTGGGGCTCCGAGCCAA 2762
QY 1740 TAAGACATCCCCACCTCTGGGGAGGGCGCGCTCCGGGGGCAATCATCCAGTGTATGG 1799
Db 2763 TAAGACATCCCCACCTCTGGGGAGGGCGCGCTCCGGGGGCAATCATCCAGTGTATGG 2822
QY 1800 CGATGACAGCTGACAGGCGGCGCAGCAGCTTATCCCTCTACTCTCTCATCTGCTCCAGCCAGC 1859
Db 2823 CGATGACAGCTGACAGGCGGCGCAGCAGCTTATCCCTCTACTCTCTCATCTGCTCCAGCCAGC 2882
QY 1860 CCAGCTATGCTTCCATGGGCAACCGGATGCGGTGAAGTTCCTTGTCTCGTGGCCAGGGAA 1919
Db 2883 CCAGCTATGCTTCCATGGGCAACCGGATGCGGTGAAGTTCCTTGTCTCGTGGCCAGGGAA 2942
QY 1920 CGTGTGCGCACCTCTGAATGCGAGTGTCTGGACAGCCAGCGAGGCGCTTGGGCGCAGC 1979
Db 2943 CGTGTGCGCACCTCTGAATGCGAGTGTCTGGACAGCCAGCGAGGCGCTTGGGCGCAGC 3002
QY 1980 TGCCCTCTGCTCGAGGTGCGAGGCGCAGAAAGCTGCGGAAAGTGTCTGCTGAGCGGCG 2039
Db 3003 TGCCCTCTGCTCGAGGTGCGAGGCGCAGAAAGCTGCGGAAAGTGTCTGCTGAGCGGCG 3062
QY 2040 GGAGGGCTACATCGACTTCGCAATTTGGAGACGAGAGGACGACGAGACGAGGAGCGCG 2099
Db 3063 GGAGGGCTACATCGACTTCGCAATTTGGAGACGAGAGGACGACGAGACGAGGAGCGCG 3122
QY 2100 AGGGACATGACCGAGGTGAGCCCGTGTCTGTCAGAGGACAGCGAGTCAATCATCTG 2159
Db 3123 AGGGACATGAGCCAGGTGAGCCCGTGTCTGTCAGAGGACAGCGAGTCAATCATCTG 3182
QY 2160 GTGGCAGGTGTCTACACCCCGAGTGAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2219
Db 3183 GTGGCAGGTGTCTACACCCCGAGTGAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3242
QY 2220 AGGACCCCGCAGACACTGACCCCGCGCGCGGGGTAGCCAGCAGCGCGCGCGCG 2279
Db 3243 AGGACCCCGCAGACACTGACCCCGCGCGCGGGGTAGCCAGCAGCGCGCGCGCGCGCG 3302
QY 2280 CCTCTTCTAACTCTCAACCTGACGCTTTCACCTGAGTGTGGCCCTCTCAGCGGCGAG 2339
Db 3303 CCTCTTCTAACTCTCAACCTGACGCTTTCACCTGAGTGTGGCCCTCTCAGCGGCGAG 3362
QY 2340 GAGTGGCGGGATGCGGATCAGCTGGAGAGGAGGAGGGGTGCTTCCACCCGAGGGGA 2399
Db 3363 GAGTGGCGGGATGCGGATCAGCTGGAGAGGAGGAGGGGTGCTTCCACCCGAGGGGA 3422
QY 2400 AGATGCTCTGCGGACGATTTTCCCGGACGCTCTGCGGACGCTTCCAGCCCGCAGAGTCTCTCA 2459

QY 678 AGAGGACC-----CGGCGCAGATGGCGTCTG 705
Db 2453 AAAGCATCTTTATGTGGAAGTATGACAAAGCAACAGCTCAGCAGAGACAGACAGCCTGTGA 2512
QY 706 GCGGTATCAACCTGGTGGGTGTC--CACCCCTGCAACGTGCCGCGGAGCAACTGCT 763
Db 2513 GGAGGCATCAGTGGTGGTGTGTCGAGAAGTGTGACGGAGTGCCTCCCT 2572
QY 764 CCTCCGAGGGGACACC--CCAGTCTAGACAAAGGGGAGGGGAGTGGCCACCATCGCC 822
Db 2573 AGTACAAATGCTGCTTCTCCAGTGTATGATATAACCAACAGAAATGGAAGCAGAAATAGT 2632
QY 823 AACGGGAAGTCAACCGTCCAGTCCACAGAGAGGCCACAGAGGCCACGAGGTGCA 882
Db 2633 GA--GGTGTATGAAATGTTCCAAACAGCAGAGAAAGCACTGAAGCTTACAGAGGG--A 2688
QY 883 GACCTCGGCGCCAGCGAGCAGAGACAGCCACATTTGCGGCCCGGCGCTCTCCACAGAGCAC 942
Db 2689 ATGCGGGTCACTGAGAGACAGATTTGACATCTCCCAAACTGGGCTCTACACAGAGCAT 2748
QY 943 GTCTTCACTGACCCAG-----CCCCGACCCCGTCTCTGGCCCCCAGCC 986
Db 2749 GTCTTTACAGATCCTTTTGGGAAGTTTCAGATCCAGAGACCTCTCCCCAGTGTATCAGTC 2808
QY 987 TGGCAGCAGAACCGGCCAGAGCCTTGACAGCAGCAGCAGCCCA--GAGCCAGAGCC 1043
Db 2809 GAGCAATGACTCAGATGCATATAAAGATCAAAATACGATCTGCCCAATGAAAGACTT 2868
QY 1044 CAGCGGGACCCACCGGAGCAGCAGAGTCTGCACCCACCATGTGGCTGGGAGCCCA 1103
Db 2869 GGTGAGAGAGNAGCCGAGAAATGAGTAGTCTTTTACCACTAATGTGGCTGGAGTCA 2928
QY 1104 GAACGGCTGGCTTATGTGCACTCGGCTGTGGCCAACTGGAAGAAGTGCCTGCATCCAT 1163
Db 2929 AAAATGGCTGTTGTATGTCATTCATCTGTAGCCAGTGGAGAAATGTCTCCATTCAT 2988
QY 1164 CAGCTGAAGATTTCTGCTGAGCCTGGTGTATCTCAAGCCCGTGTGCTGCTGCTCT 1223
Db 2989 TAAACTTAAAGATTCGATTTCTCAGTATTGTACACGTGAAGGAAATCGTTAGTAGCCCT 3048
QY 1224 GCGCAGCGGACCTTGGCCATCTCCACGCTGGTGAAGATGCCAGTGGGATCTGAGCAA 1283
Db 3049 GCTCAGCGCACCTTGCATCTTTTCAGAGAGTGGATGGCAGTGGGATTTGTCAA 3108
QY 1284 CTATCACTTAATGGAACCTGGGCCACCGCACCTCCATCCGCTGATGCTGTGTGTA 1343
Db 3109 CTATCACTTTAGACCTTGGACGGCTCATCATCTCCATCCGTTGCACTGTGTACA 3168
QY 1344 CGACCGCTGTGTGTGCTACAGAACAGGTGCAAGTGCACGTCATCCAGCCCAAGACCATGA 1403
Db 3169 TGACAAAGTCTGGTGTGGCTATAGGAACAAATCTATGTGGTGACCCAAAGGCCATGA 3228
QY 1404 GATAGAGAGTCAATTTGACGCCACCCCGGGGGAGAGCCAGGTGCGGAGTGGCGTG 1463
Db 3229 AATAGAGAAATCTTTTGTATGCACTCCAGGAGGAGAGCCAGTGGCAGAGTTCGCTG 3288
QY 1464 GATCGCGATGCGTATGGTGTCCATCCGCTGAGCTCCACCTGAGGCTCTACCATGC 1523
Db 3289 GGTGGGGATGCGGTGTGGCTCCCATTCGCTGGATTTCTACGCTCCGCTCTCTATCATGC 3348
QY 1524 ACACCGCACCATCTACAGGACGTGGACATTTGAGCCCTACGTGAGCAAGATGCTAGG 1583
Db 3349 ACACACTTATCACTATCAGAGATGTGACATTTGAGCCCTTATGTAGCAAAATGTTAGG 3408
QY 1584 CACTGCAAGCTGGGTTTCTCTTCTGATGATCAGGCCCTCTTGTGCGGGCAGCCG 1643
Db 3409 TACTGGAACACTGGGCTTCTCTTTGTGAGAAATACAGCTCTTATGTGTCTTGTAAATCG 3468
QY 1644 GCTCTGGGTGGGCACCGGCACAGGATGTGCTCTCCATCCCTGACAGAGCTGTGT 1703
Db 3469 TTTGTGGGTGGGACAGGAAATGGTGTCTTCTCATTCATCCCATTTGACAGAAACCGTAAT 3528
QY 1704 CCTGCACCGAGGCCAGCTCTCTGGGGCTCCGAGCCAAATAGACATCCCCCACCTCTGGGGA 1763

Db 3529 CTTCCACAGAGGAGCTTTTACTGGGCTGAGGCGAAATAAAACCTTCAGGTGTACAGAAA 3588
QY 1764 GGGGCGCCGTCCTCCGGGGCATCATCCACGTGTATGGGATGACAGCAGTGTGACA--GGGC 1820
Db 3589 T-----CGTCCTGGAAAGTGTAAATCCGTGTATATGGTGTATGAAACAGTGTATAAGTGAC 3642
QY 1821 GGCCAGCAGCTTCATCCCTTACTCTCATGGCCCCAGGCCAGCTATGCTTCCATGGCA 1880
Db 3643 TCCAGGACATTTATACCTTATGTTCAATGGCACATGCACAGCTTTGCTTCCATGGCA 3702
QY 1881 CCGCATGCCGTGAAGTCTTTTGTCTGGTCCAGGGAACGTGTGCGCCACCTGAAATGG 1940
Db 3703 CCGGATGCTGTGAAATTTCTTTGTGGCAGTCCCAGGTCAAGTCAATCAGCCCAAAAGTAG 3762
QY 1941 CAGTGTCTGACAGCCAGCCAGGCGCCCTGGGCCAGCTCCCTGCTCGAGGTGCA 2000
Db 3763 CAGT--AGTGGCACCGATCTGACGGGTGACAAAGCAGGGGCCATCTGCACAGGAGCCTGG 3821
QY 2001 GGGCCAGAAGC---TGCGGAACGTGCTGGTCTGAGCGCGGGAGGCTACATCGACTT 2057
Db 3822 TAGTCAGAGCCCTTGAAGTCTATGCTTGTATCAGTGGAGAGAGGGCTACATCGACTT 3881
QY 2058 CCGCATTTGACAGCGAGAGAGCAGCAGCAGCAGGAGAGGGCCGAGGACATGAGCCAGT 2117
Db 3882 CCGAATGGGTGATGAAG---GTGGGAATCAGAACTTCTTGGAGAGGATCTTCCACTT 3936
QY 2118 GAAGCCCGTGTCTCCAGGAGCAGCAGTCACATCATCTGTGCGAGGTGTCTTA 2174
Db 3937 GAACCTTCTG--TCACCAAGCAGAAAGAGTCACTTGTAGTGTGGCAAGTATGTA 3992

RESULT 3

US-10-044-090-549
; Sequence 549, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:

; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 549
; LENGTH: 4286
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1327121.3
US-10-044-090-549

Query Match 12.9%; Score 477; DB 14; Length 4286;
Best Local Similarity 56.5%; Pred. No. 5.5e-110;
Matches 1154; Conservative 0; Mismatches 805; Indels 84; Gaps 11;

QY 214 GSCCAGGAGGACACCGGATGAAGACGTGCGCGGTGCTACTCGCGCCCTCTGCTG 273
Db 1030 GGTCAAGGTGAAATTAAGATGAAATTTACCTGTGCTCTATCTCAGACTCTGGAT 1089
QY 274 GAGAAGGACCCACCATGAAGCTGTGTGCGCGCGGCTCAACCTGAGCGGGTGGAGG 333
Db 1090 GAAAGATACATCAATGAAGCTGTGTGCTGTGTTGGTCAATTTATCTGTGGGAAG 1149
QY 334 CCAATAGAGACGACCTGGGAATGAG-----TCNAGCCAGCCAGCCCGC 381
Db 1150 ACCAGAGATGGTGGTCTCTGTTGTGGCAAGTGTATTTACAAGGATGTGCTGTTG 1209
QY 382 GATCCCTGACCTGCCACCGGAGAGGAGCGGAGCCCAAGAGCCGCCACAGCTCTCCC 441
Db 1210 GATACAGAGGCGAGTAAACAGCGAAGTGCCTCTCAGAGTAGTTTAGATAGTATCAG 1269

QY	442	GAG---	AGAAGAAGCCAAAGAGCTCCCTGAAATGGAGCCACCCTCCAGCCGGGTGGG	499
Db	1270	GAACTTAAGGAACAGCAGAAGGAGTTAAAAAATCAAGAAGAAATATCCAGTCTAGTTTGG	1329	
QY	499	ATCCCTGACCAGCACCTGACACACAGCAAGGTGGTGATCATCGAGCCAAACAGCCCGGC	558	
Db	1330	ATCTGTACAGCACTCATTCGGCTACAAAGTTCCTATTATGATGCTGTTCAACCTGGC	1389	
QY	559	ACGGTGGTGGACCAAGTTCACCGTCTGCAACGGCAGCTGCTGCATCTCCAGCATCCCC	618	
Db	1390	AAACATCCTTAGACAGTTTCACTGTTTGGCACTCTCATGTTCTGTGATGCAAGTGTGCCA	1449	
QY	619	GCGGCAGGACAGCGACTACCCCTCCGG-GGAGATGTTCTTGGACAGCAGCGTGAACCC	677	
Db	1450	GGTGACAGAAACAGACTTACCCTGCAGAGAAGATCTTTCAGAATCTGGTCAGTAGAC	1509	
QY	678	AGAGAGCC-----	705	
Db	1510	AAAGCATCTTTATGTGGAGTATGACAAGCAACAGCTCAGCAGACAGACAGCCTGTTA	1569	
QY	706	GCCGTATCACCCCTGGTGGGCTGTGCCACCCGCTGCAAGCTGCGCGAGCACTGCTCC	765	
Db	1570	GGAGCATCACAGTGTGTTGTTCTGCAGAAAGTGTGACGGAGCTGCCACTTCCCT	1629	
QY	766	TCCCGAGGGGACACCCCAAGTCTAGACAAGGGGAGGGGAGGTGGCCACCATCGCCAAC	825	
Db	1630	AGTACAAATGGTGTCTTCCAGTGATGATATAAACCAACAGAAATGGAAGCAGAAATAGT	1689	
QY	826	GGGAAGGTCAACCCGTCCCACTCCACAGAGGAGGCGCACAGAGGCCAGGAGTGCAGAC	885	
Db	1690	GAGTTGATGAAATGTTCCMACAGCAGAGAAGCAACTGAAGCTACAGAAGGGA--AT	1746	
QY	886	CCTGGGCCACGAGCCAGACAGCACACATTTGGGCGCGGGCTCTTCACAGAGCAGTCC	945	
Db	1747	GCGGGTCAGCTGAAGACACAGTGGACATCTCCCAAACTGGCGTCTACACAGAGCATGC	1806	
QY	946	TTCACTGACC-----	990	
Db	1807	TTTACAGATCCTTTGGAGTTTCAGATCCCAAGAGACCTCTCCCAGTGTATCAGTCGAGC	1866	
QY	991	AGCAGAAACGGCCAGAGCCTGACAGCAGCAGCACACGGCCA--GAGCCAGAGCCCAAGC	1047	
Db	1867	AATGACTCAGATGATATAAGATCAAAATACGACTGCCAAATGAAACAGACTTTGGTG	1926	
QY	1048	GGGACCCCCACGGAGCAGCAGCAGTGTGCAACCCACCATGTGGCTGGAGCCCAAGAC	1107	
Db	1927	AGAAAGAACGCCAGAAAAATGAGTAGTCTTTTACCAACTATGTGGCTTGGAGCTCAAAAT	1986	
QY	1108	GGCTGGCTCTATGTGCACTCGGCTGTGGCCAACTGGAAAGATGCTGCACTCCATCAAG	1167	
Db	1987	GGCTGTTGTATGTCCATTCATCTGTAGCCCACTGGAGGAAATGTCTCCATTTCCATTAAA	2046	
QY	1168	CTGAAGATTTCTGTGTAGCCTGTGGTGCATGTCAAAGGCCGTGTGCTGTGGTCTTGGG	1227	
Db	2047	CTTAAAGATTCGATTTCTCAGTATTTGACGTOAGGGAAATCGTGTATTAGCCCTGGCT	2106	
QY	1228	GACGGACCCCTGGCCATTTTCCACCGTGGTGAAGATGGCCAGTGGGATCTGAGCAACTAT	1287	
Db	2107	GACGGCACCCCTGCAATCTTTACAGAGAGTGGATGGCAGTGGGATTTGTCAAACATAT	2166	
QY	1288	CACCTAATGCACCTGGCCACCGGCAACACTCAATCCGCTGATGGCTGTGTGTACGAC	1347	
Db	2167	CACCTCTTAGACCTTGGACGGCTCTCATCTCCATCCGTTGCATGACTGTGTGTCATGAC	2226	
QY	1348	CGCGTGTGGTGTGGCTACAGAACAGGTGCAGTCACTCAGCCCAAGACCATGCAATA	1407	
Db	2227	AAAGTCTGTGTGGCTTATAGAAACAAAATCTATGTGTGTGACCCAAAGGCCATTGAAAAATA	2286	
QY	1408	GAGAAGTCAATTTGACGCCCAACCCGCGCGGGAGAGCCAGTTCGGCAGCTGCGTGGATC	1467	
Db	2287	GAGAAATCTTTTGATGCACATCCAGAAAGAGAGCCAGTTCGCAAGCTTTCGCTGGGTG	2346	
QY	1468	GGCGATGGCGTATGGGTGTCCATCCGCTTGGACTCCACCTTGAGGCTCTACCATGCAAC	1527	

Db	2347	GGGGATGCGGTGGGTCTCCATTGCTTGGATTCTACGCTCCGCTCTCTATCATGCACAC	2406
QY	1528	ACGCACCAAGCATCTACAGGAGCTGGACATTAGCCCTACGCTCAGCAGATGCTAGGCACT	1587
Db	2407	ACTTATCAACATCTACAGGATGGACATTGAGCCCTTATGTAAGCAAAATGTTAGSTACT	2466
QY	1588	GGCAAGCTGGGTTTCTCTTGTAAGCATCACGGCCCTGCTGTCGGGCAGCCGGCTC	1647
Db	2467	GGAAAACTGGCTTCTCTTTGTGAGAAATTACAGCTCTTATGCTGCTTGTAATCGTTTG	2526
QY	1648	TGGGTGGGCACCGCAACGAGTGGTTCATCTCCATCCCCTGACAGAGACTGTGGTCCTG	1707
Db	2527	TGGGTGGGACAGGAATGGTGTCAATTATCTCCATCCCATTTGACAGAAACCGTAATCCTC	2586
QY	1708	CACCGAGCCAGCTCTCTGGGGCTCCGAGGCCAATAAGACATCCCCCCTCTGGGGAGGGC	1767
Db	2587	CACCAGGAGCGTTTACTGGGCTCAGGGCAATAAACCCTCAGGTGTACAGGAAT---	2643
QY	1768	GCCCGTCCCGGGGCATCATCAGTGTATGGCGATGACAGCAGTGAACA---GGCGGGC	1824
Db	2644	---CGTCTCGAAGTGTAAATCCGTGTATATGTTGATGAAACAGTCAAAAGTGACTCCA	2700
QY	1825	AGCAGCTTCATCCCCTACTCTCCATGGCCCGCCAGCCAGCTATGTTCCATGGCACCCG	1884
Db	2701	GGACATTTATACCTATTTGTTCAATGGCACATGCACAGCTTGTCTTCCATGGCACCCG	2760
QY	1885	GATGCCGTGAAGTCTTTGTCGGTGCCAGGGAACGTGTGCGCCACCCCTGTAATGSCAGT	1944
Db	2761	GATGCTGTGAAATCTTTGTGGCAGTCCCAGGTCAGTCAATCAGCCCAAAAGTAGCAGT	2820
QY	1945	GTGCTGACAGCCAGCCGAGGCCCTGGCCAGCTGCCCTTGCCCTCGAGGTCGAGGGC	2004
Db	2821	AGTGGCAGCGATCTGACGGGTGACAAAGCAGGGCCATCTGCACAGGAGCTGTAAGTCAG	2880
QY	2005	CAGAAGCTGGGAACGTGTGTTGCTGACGGCGGGAGGGCTACATCGACTTCGCAATT	2064
Db	2881	ACGCCCTTGAAGTCTATGCTTGTTCATCACTGAGAGAGAGGGCTACATCGACTTCGCAATG	2940
QY	2065	GGAGACGGAGAGACACAGAGACGGAGGGCGCAGGGGACATAGGCCAGGTGAAGGCC	2124
Db	2941	GGTGATGAAG-----GTGGAGAAATCAGAACTTCTTTGGAGAGATCTTCCACTTTGAACCTT	2995
QY	2125	GTGCTCTCAAGGACAGGCGCAGTCACATCATCGTGTGGCAGGTGTCTACACCCCGAG	2184
Db	2996	CTG-TCACCAAGACGAAAGGAGTCACTTGATGTGGCAAGTATGATGGCAATGAG	3054
QY	2185	TGA	2187
Db	3055	TGA	3057

```

RESULT 4
US-10-240-965-3
; Sequence 3, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, Dov
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAWN, Richard M.
; APPLICANT: SEILHAMER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; APPLICANT: TAI, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program

```

; SEQ ID NO 3		1048	GGGACCCACCGGAGCAGGCGAGCTGCTGACCCACCATGTGGCTGGAGCCGACGAAAC	1107
; LENGTH: 3791				
; TYPE: DNA		Db		
; ORGANISM: Homo sapiens		2944	AGAGAGAGCCAGAAAATGAGTAGTCTTTTACCAACTATGTGGCTTGGAGCTCAAAAT	3003
; FEATURE:				
; NAME/KEY: misc feature		1108	GGCTGGCTCTATGTGCACTCGGCTGTGGCCAACTGGAGAAAGTGCCTGCACTCCATCAAG	1167
; OTHER INFORMATION: Incyte ID No. US20030165924A1 247178.2				
US-10-240-965-3		3004	GGCTGTTTGTATGTCCATTCTCTGTAGCCAGTGGAGGAATGTCTCCATTCATTAATA	3063
Query Match		1168	CTGAGGATCTGTGCTGAGCCCTGCTCATGTCAAGCCGCTGTGCTGGTCTGGCG	1227
Best Local Similarity			9.8%; Score 360.6; DB 15; Length 3791;	
Matches		991; Conservative	0; Mismatches 669; Indels 119; Gaps 11;	
		214	GGCAGGAGGACACGCGATGAAGAAAGTGGCGGTGCGGGTACTGCGGCCCTCTGGTG	273
		2047	GGTCAAGGTGAAATAAAGATGAAAAATTACCTGTGCTGTCTATCTCAGACCTCTGGAT	2106
		274	GAGAGGACCCACCATGAAGCTGTGTGTCGCGGGCGTCAACCTGAGCGGTGGAGG	333
		2107	GAAAAAGATACATCAATGAAGCTGTGTGCTGTGTGGAGTCAATTTATCTGTGGGAG	2166
		334	CCCAATGAGGACGACGCTGGGAATGGAG-----TCAAGCCAGCGCCAGCGCGC	381
		2167	ACCAGATGGTGTCTGTGTTGGACAAAGTGTATTTTACAGGATGTTGCTGGTTG	2226
		382	GATCCCTGACCTGCGACCGGAAGAGAGCGCGAGCCCAAGAGCGGCCACAGCTCTCCC	441
		2227	GATACAGAAAGGAGTAAACAGCGAAGTGCCTCTCAGAGTAGTTTAGATAAGTTAGATCAG	2286
		442	GAG---AAGAAGAGCCCAAGAGCTCCTGAAATGGAGCGCACCTCCAGCGGGTGTGG	498
		2287	GAACTTAAGGAACACAGAGAAGAGTTAAAAATCAAGAGAATTAATCAGTCTAGTTGG	2346
		499	ATCTGTACACGACCTGACCAACAGCAAGGTGTGATCATCGACCGCCCAACAGCGGGC	558
		2347	ATCTGTACACGACCTCATTCGGCTACAAAGTTCTTATTAATGATGCTGTTCAACTGGC	2406
		559	ACGGTGGTGGACAGTTACCGTTACCGTGAACGGCGAGTGTGTGATCTTCAGCATCCCC	618
		2407	AACTCTTAGACAGTTTCACTGTTGCAACTCTCATGTTCTGTGATTTGCAAGTGTGCA	2466
		619	GGGGCAGCGACAGCACTACCTCCCGG-GGAGATGTTCTCTGGACAGCGAGCTGAACCC	677
		2467	GGTGACAGGAACAGACTACCTCTGAGGAGAGATCTTTTCAAGATCTGCTCAGTAGAC	2526
		678	AGAGAGACC-----CGGGCGCAGATGGCGTGTG	705
		2527	AAAGCATCTTATGTGGAAGTATGACAAGCAACAGCTCAGCAGAGACAGACAGCGCTGTTA	2586
		706	GCCGCTATCACCTGTGGGTGTGCGCACCGCTGCAACGTGCGCGGAGCAACTGCTCC	765
		2587	GGAGGATCACAGTGTGTGTTCTTGAGAGAGTGTGACGGGAGCTGCCACTTCCCT	2646
		766	TCCGAGGGACACCCAGTGTGACAAAGGGGAGGTGGCCACCATCGCCAAC	825
		2647	AGTACAAATGGTCTTCTCCAGTGTGATGAATAAACACAGAAATGAAGCAGAAATAGT	2706
		826	GGGAAGGTCAACCCGCTCCAGTCCACAGAGGAGGCCACAGAGCCACGAGGTGCGAGAC	885
		2707	GAGTTGATCAAAATGTTTCAACAGCAGAGAAGCAACTGAAGCTACAGAAGGG---AAT	2763
		886	CTGGGCCAGCGAGCCAGAGACAGCCCATTTGCGCGCGGGCTCTCACAGAGCAGCTC	945
		2764	CGCGGTCTCAGCTGAAGACACAGTGGACATCTCCAAACTGGCGCTTACACAGAGCATGC	2823
		946	TTCACTGACCC-----AGCCCCGACCCCGTCTCTGCGCCCGCCAGCCTGGC	990
		2824	TTTACAGATCTTTGGAGTTTCAGATCCCGAGAGACCTCTCCCGAGTGTATCAGTCGAGC	2883
		991	AGCGAGAACGGGCGAGAGCCCTGACAGCAGCAGCAGCGGCCA---GAGCGAGCCGACG	1047
		2884	AATGACTCAGATGCATATAAGATCAAAATATCAGTACTGTCGCAATGAACAGACTTTGGT	2943
		Db		

RESULT 5
US-09-925-297-167
; Sequence 167, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08


```

; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 167
; LENGTH: 1413
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1376)
; OTHER INFORMATION: n equals a.t.g. c
US-09-925-297-167

```

Query Match.	6.6%;	Score 242.4;	DB 9;	Length 1413;
Best Local Similarity	62.5%;	Pred. No. 5.8e-51;		
Matches 451;	Conservative 0;	Mismatches 256;	Indels 15;	Gaps 4;
QY	1469	GC	GATGCGCTATGGGTGCTCCATCGCTGCAGCTCCACCCCTGAGGCTTACATGACACACA	1528
Db	1	GG	ATGCGGTGGGTCTCCATTCGCTTGGATTCTACGCTCCGCTCTCTATCATGACACACA	60
QY	1529	CG	CACAGCATCTACAGGACGCTGACATGAGCCCTACGTCAGCAAGATGCTAGGCACTG	1588
Db	61	CT	TATCAACATCTACAGGATGGACATTGAGCCTTATGTAAGCAAAATGTTAGGTACTG	120
QY	1589	GC	AGCTGGGTTTCTCTTCGTAGCATACGGCCCTGCTTGTGCGGGGAGCCGCTCT	1648
Db	121	GA	AACTGGGCTCTCTTTTGTGAGATTACAGCTCTTATGTTGTCTTGTAAATCGTTTGT	180
QY	1649	GG	TGGGACCGGCAACGGAGTGGTCACTCCATCCGCTGACAGAGACGTGGTCTCTGC	1708
Db	181	GG	TGGGACAGAAATGGTGTCTATTATCTCCATCCCATGACAGAAACCGTAATCCTCC	240
QY	1709	AC	CGGCGAGCTCTGGGGCTCCGAGCCAAATAAGACATCCCCCAGCTCTGGGGAGGGCG	1768
Db	241	AC	CAGGAGCTTTACTGGGGCTGAGGGCAATAAATA-----CCTCAGGTGTACAGGAA	294
QY	1769	CC	CGTCCCGGGGCATCCACGTGTATGGCGATGACAGCAGTGACA---GGGCGGCCA	1825
Db	295	AT	CGTCTGGAAGTGAATCCGTGTATGTTGATGATAAACAAGTGATGACTCCAG	354
QY	1826	GC	AGCTTCATCCCTACTGCTCCATGCCCGAGCCAGAGTATGCTTCCATGGGACCGGG	1885
Db	355	GG	ACATTTATACCTATGTTCAATGGCACATGCAAGCTTTGCTTCCATGGGACCGGG	414
QY	1886	AT	GCGTGAAGTCTTTTGTCTGGTGCAGGGAAACGTGTGGCACCTGAAATGGCAGTG	1945
Db	415	AT	GCTGTAATCTTTTGTGGCAGTCCCAGGTCAAGTCAATCAGCCACAAAGTAGCAGTA	474
QY	1946	TG	CTGACAGCCAGCGGAGGGCCCTGGGCCAGCTGCCCTGCCTCGAGGCTCGAGGGCC	2005
Db	475	GT	GCAAGGATCTGACGGGTGACAAAGAGGGCCATCTGCAAGGAGCCTGGTAGTCAGA	534
QY	2006	AGA	GCTCGGAAACGTTGCTGGTCTGAGCGGGGAGGGGTACATCGACTTCCCGATTG	2065
Db	535	CG	CCCTTGAAGTCTATGCTTGTTCATCAGTGGAGGAGGGCTACATCGACTTCCGAAATGG	594
QY	2066	GAGA	CGGAGGACGACGAGACGGAGAGGGCGAGGGGACATGAGCCAGGTGAAGCCCG	2125
Db	595	GT	GATGAAG-----GTGGAGAATCAGAACTTCTTGAGAGGATCTTCCACTTGAACCTTC	649
QY	2126	TG	CTGTCCAAGGACAGGCCAGTCACATCATCGTGTGGCAGGTGTCTTACACCCCGAGT	2185
Db	650	TG	-TCACCAAGACGAGAAAGGATCACTTCGATAGTGTGGCAAGTGTATGTCGCAATGAGT	708
QY	2186	GA	2187	
Db	709	GA	710	

RESULT 6
US-10-085-783A-1790

```

; Sequence 1790, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1790
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-1790

```

Query Match.	6.5%;	Score 241.8;	DB 13;	Length 310;
Best Local Similarity	94.0%;	Pred. No. 6.8e-51;		
Matches 296;	Conservative	0;	Mismatches 12;	Indels 7;
				Gaps 4;
QY	2824	CTGGGGCAGCTGCTTGAGACAGAGATGCTACCCCATCTCTGCCATGCGAGGCAGGCTCT	28833	
Db	1	CTGGGGCAGCTGCTTGAGAAACAGAGATGTAACCCATCTCTGCCCATGCGAGGCAGGCTCT	60	
QY	2884	TGCCAGCCCCGTTCTGACCCCGTGTCCCCCGAGGCTCTGCTTGGGCGAGAGACTCACCTTG	2943	
Db	61	TGCCAGCCCCGTTCTGACCCGCTGTCCCCCGAGGCTCTGCTTGGCGAGAGACTCACCTTG	120	
QY	2944	GAGAGTGGGCCCTT-GGAGTCTCTGCCCTCCCAAGAGCCCCCAGGGT-GGGATTTCTCAG	3001	
Db	121	GAGGAGTGGGCCCTCTGGAGTGCTGTCCCTCCCAAGAGCCCCCAGGGTGGGGATTTCTCAG	180	
QY	3002	GCTGCCAGGGCAGGCCCGCCAGGCTCAGAGAGAGGGAGGGCCCTTGGCCTCTCCGGGATCA	3061	
Db	181	GCCTGTCAGTGAGGGCCCGAGGCTCTCAGAGAGAGGGAGG--CCCTGGCTCTCCGGGATCA	237	
QY	3062	GTCTTAGGACACAGGCTCAGCCCTCAGGTTGATGGGGGATGATGTGTCCCGGGGCTGCC	3121	
Db	238	GTCTTAGGACACAGGCTCAGCCCTCAGGTTGATGGGGGATGATGTGTCCCGGG--CTGCC	295	
QY	3122	TCTTGCAACGGGGCTC	3136	
Db	296	TCTTGCAACGGGGCTC	310	

RESULT 7
US-10-242-535A-1790
Sequence 1790, Application US/10242535A
Publication No. US20040013663A1
GENERAL INFORMATION:
APPLICANT: ChondroGene Inc.
APPLICANT: Liew, C.C.
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
FILE REFERENCE: 4231/2005
CURRENT APPLICATION NUMBER: US/10/242,535A
PRIOR FILING DATE: 2002-09-12
PRIORITY APPLICATION NUMBER: US 10/085,783
PRIOR FILING DATE: 2002-02-28
PRIORITY APPLICATION NUMBER: US 60/305,340
PRIOR FILING DATE: 2001-07-13
PRIORITY APPLICATION NUMBER: US 60/275,017
PRIOR FILING DATE: 2001-03-12
PRIORITY APPLICATION NUMBER: US 60/271,955
PRIOR FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 58994
SOFTWARE: PatentIn version 3.2

```
; SEQ ID NO 1790
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-1790

Query Match
Best Local Similarity 6.1%; Score 241.8; DB 16; Length 310;
Matches 296; Conservative 0; Mismatches 12; Indels 7; Gaps 4;

QY 2824 CTGGGGGAGCTGCTTTGAGAACAGAGACTGCTACCCCACTCCCTGCCCATGCGAGGAGGCTCT 2883
Db 1 CTGGGGGAGCTGCTTTGAGAACAGAGACTGCTACCCCACTCCCTGCCCATGCGAGGAGGCTCT 60
QY 2884 TGCCAGGCCCCGTTCTGACCCGCTGTCCTCCAGGCTCTGCTGGGAGAGACTCACCTTG 2943
Db 61 TGCCAGGCCCCGTTCTGACCCGCTGTCCTCCAGGCTCTGCTGGGAGAGACTCACCTTG 120
QY 2944 GAGAGTGGGCGCTTGGAGTCTCTGCTCCCTCCAGAGCCCCAGGCTTGGGATTTCTCAG 3001
Db 121 GAGAGTGGGCGCTTGGAGTCTCTGCTCCCTCCAGAGCCCCAGGCTTGGGATTTCTCAG 180
QY 3002 GCTGCCAGGCGAGGCCAGGCTCTAGGAGAGAGGAGGCGGCTGCTCCGAGTCA 3061
Db 181 GCTGCCAGGCGAGGCCAGGCTCTAGGAGAGAGGAGGCGGCTGCTCCGAGTCA 237
QY 3062 GTCTTAGGACACAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 3121
Db 238 GTCTTAGGACACAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 295
QY 3122 TCCTGCACGGGGCTC 3136
Db 296 TCCTGCACGGGGCTC 310

RESULT 8
US-10-264-237-807
; Sequence 807, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; APPLICANT: PA131PI
; FILE REFERENCE: US/10/264,237
; CURRENT APPLICATION NUMBER: US/10/264,237
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 807
; LENGTH: 554
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (361)..(361)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (416)..(416)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-807

Query Match
Best Local Similarity 6.1%; Score 226; DB 16; Length 554;
Matches 373; Conservative 1; Mismatches 137; Indels 6; Gaps 6;

QY 1069 AGCAGTGTGACCCACCTGCTGGCTGGGAGGCCAGGCTGCTATGTGCACTCG 1128
Db 29 AGTAGTCTTTTACCAACTATGCTTGGCTTGGAGCTCAAAATGGCTGTTTGTATGTCCATCA 88

; SEQ ID NO 965
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:

; Sequence 965, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 4921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 965
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
```

```

; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 6852
; LENGTH: 189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-6852

Query Match          5.1%; Score 189; DB 9; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.3e-37;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2833  CTGCTTGAGAACAGAGACTGCTACCCCATCTGCCCATGCAGGCAGGCTTTGCCAGCCC 2892
Db      1      CTGCTTGAGAACAGAGACTGCTACCCCATCTGCCCATGCAGGCAGGCTTTGCCAGCCC 60

Qy      2893  CGTCTTGACCCGTGTCCCCCAGGCTCTGCTGGGAGAGACTCACTTGGAGGAGTGG 2952
Db      61     CGTCTTGACCCGTGTCCCCCAGGCTCTGCTGGGAGAGACTCACTTGGAGGAGTGG 120

Qy      2953  GCCTTGAGTCTGTCCCTCCCAAGAGCCGCCAGGCTGGGATTTCTCAGGCTGCCAGGCC 3012
Db      121    GCCTTGAGTCTGTCCCTCCCAAGAGCCGCCAGGCTGGGATTTCTCAGGCTGCCAGGCC 180

Qy      3013  AGGCCCCAGG 3021
Db      181    AGGCCCCAGG 189

RESULT 11
US-10-040-862-6852
; Sequence 6852, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545

```

```

; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6852
; LENGTH: 189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-6852

Query Match          5.1%; Score 189; DB 15; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.3e-37;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2833 CTGCTTGAGAACAGAGACTGTACCCCATCTGCGCCATGCGAGGAGGCTTTGCCAGGCC 2892
Db 1 CTGCTTGAGAACAGAGACTGTACCCCATCTGCGCCATGCGAGGAGGCTTTGCCAGGCC 60

QY 2893 CGTTCTGACCCGTGTCCCCCAGGCTCTGCTGGCGCAGAGACTCACCTTGAGGAGTGG 2952
Db 61 CGTTCTGACCCGTGTCCCCCAGGCTCTGCTGGCGCAGAGACTCACCTTGAGGAGTGG 120

QY 2953 GCCTTGGAGTCTCTGCTCCCTCCAGAGAGCCCGCCAGGCTGGGATTTCTCAGGCTGCCAGGCG 3012
Db 121 GCCTTGGAGTCTCTGCTCCCTCCAGAGAGCCCGCCAGGCTGGGATTTCTCAGGCTGCCAGGCG 180

QY 3013 AGGCCCCAGG 3021
Db 181 AGGCCCCAGG 189

RESULT 12
US-10-057-475B-6852
; Sequence 6852, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17

```

```

; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6852
; LENGTH: 189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-6852

Query Match          5.1%; Score 189; DB 16; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.3e-37;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2833 CTGCTTGAGAACAGAGACTGTACCCCATCTGCGCCATGCGAGGAGGCTTTGCCAGGCC 2892
Db 1 CTGCTTGAGAACAGAGACTGTACCCCATCTGCGCCATGCGAGGAGGCTTTGCCAGGCC 60

QY 2893 CGTTCTGACCCGTGTCCCCCAGGCTCTGCTGGCGCAGAGACTCACCTTGAGGAGTGG 2952
Db 61 CGTTCTGACCCGTGTCCCCCAGGCTCTGCTGGCGCAGAGACTCACCTTGAGGAGTGG 120

QY 2953 GCCTTGGAGTCTCTGCTCCCTCCAGAGAGCCCGCCAGGCTGGGATTTCTCAGGCTGCCAGGCG 3012
Db 121 GCCTTGGAGTCTCTGCTCCCTCCAGAGAGCCCGCCAGGCTGGGATTTCTCAGGCTGCCAGGCG 180

QY 3013 AGGCCCCAGG 3021
Db 181 AGGCCCCAGG 189

RESULT 13
US-10-154-884B-6852
; Sequence 6852, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-01352US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01

```

/ PRIOR APPLICATION NUMBER: US 60/202,084
 / PRIOR FILING DATE: 2000-05-04
 / PRIOR APPLICATION NUMBER: US 60/206,201
 / PRIOR FILING DATE: 2000-05-22
 / PRIOR APPLICATION NUMBER: US 60/218,950
 / PRIOR FILING DATE: 2000-07-14
 / PRIOR APPLICATION NUMBER: US 60/222,903
 / PRIOR FILING DATE: 2000-08-03
 / Remaining Prior Application data removed - See File Wrapper or PALM.
 / NUMBER OF SEQ ID NOS: 11290
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 6852
 / LENGTH: 189
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 US-10-154-884B-6852

Query Match 5.1%; Score 189; DB 16; Length 189;
 Best Local Similarity 100.0%; Pred.No. 1.3e-37;
 Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2833 CTGCTTGAGAACACAGAGCTGTACCCCATCTGCCATGCGAGGCGAGGCTCTTGCAGGCC 2892
 Db 1 CTGCTTGAGAACACAGAGCTGTACCCCATCTGCCATGCGAGGCGAGGCTCTTGCAGGCC 60
 QY 2893 CGTTCTGACCGTGTCCCGCCAGGCTGTGCTTGGGCGAGAAAGTCACTTGGAGGAGTGG 2952
 Db 61 CGTTCTGACCGTGTCCCGCCAGGCTGTGCTTGGGCGAGAAAGTCACTTGGAGGAGTGG 120
 QY 2953 GCCCTGGAGTCTGTCTCCCTCCAGAGGCCCGCCAGGGTGGGATTTCTCAGGCTGCCAGGGC 3012
 Db 121 GCCCTGGAGTCTGTCTCCCTCCAGAGGCCCGCCAGGGTGGGATTTCTCAGGCTGCCAGGGC 180
 QY 3013 AGGCCCCAGG 3021
 Db 181 AGGCCCCAGG 189

RESULT 14
 US-09-954-456-819
 / Sequence 819, Application US/09954456
 / Patent No. US20020115057A1
 / GENERAL INFORMATION:
 / APPLICANT: Young, Paul
 / TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
 / TITLE OF INVENTION: Sets
 / FILE REFERENCE: 689230-76
 / CURRENT APPLICATION NUMBER: US/09/954,456
 / CURRENT FILING DATE: 2001-09-18
 / PRIOR APPLICATION NUMBER: US/60/233,617
 / PRIOR FILING DATE: 2000-09-18
 / PRIOR APPLICATION NUMBER: US/60/234,052
 / PRIOR FILING DATE: 2000-09-20
 / PRIOR APPLICATION NUMBER: US/60/234,923
 / PRIOR FILING DATE: 2000-09-25
 / PRIOR APPLICATION NUMBER: US/60/235,134
 / PRIOR FILING DATE: 2000-09-25
 / PRIOR APPLICATION NUMBER: US/60/235,637
 / PRIOR FILING DATE: 2000-09-26
 / PRIOR APPLICATION NUMBER: US/60/235,638
 / PRIOR FILING DATE: 2000-09-26
 / PRIOR APPLICATION NUMBER: US/60/235,711
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,720
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,840
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,863
 / PRIOR FILING DATE: 2000-09-27
 / NUMBER OF SEQ ID NOS: 2276
 / SOFTWARE: PatentIn version 3.0
 / SEQ ID NO 819
 / LENGTH: 6289

us-10-019-495-26.rnpb

Search completed: August 24, 2004, 13:48:18
Job time : 3797 secs